CyanoGate

Short Title: A Golden Gate-based toolkit for cyanobacteria

Corresponding author: Alistair McCormick, School of Biological Sciences, University of Edinburgh, +44 (0)1316505316, alistair.mccormick@ed.ac.uk

Title: **CyanoGate: A modular cloning suite for engineering cyanobacteria based on the plant MoClo syntax**

Authors: Ravendran Vasudevan[1,2], Grant A.R. Gale[1,2,3], Alejandra A. Schiavon[1,2], Anton Puzorjov[1,2], John Malin[1,2], Michael D. Gillespie[4], Konstantinos Vavitsas[5,6], Valentin Zulkower[2], Baojun Wang[2,3], Christopher J. Howe[7], David J. Lea-Smith[4], Alistair J. McCormick*[1,2]

[1]Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, EH9 3BF, UK.
[2]Centre for Synthetic and Systems Biology, University of Edinburgh, EH9 3BF, UK.
[3]Institute of Quantitative Biology, Biochemistry and Biotechnology, School of Biological Sciences, University of Edinburgh, EH9 3FF, UK.
[4]School of Biological Sciences, University of East Anglia, Norwich, NR4 7TJ, UK.
[5]Australian Institute of Bioengineering and Nanotechnology, The University of Queensland, Brisbane, Queensland 4072, Australia
[6]CSIRO, Synthetic Biology Future Science Platform, Brisbane, Queensland, 4001, Australia.
[7]Department of Biochemistry, University of Cambridge, CB2 1QW, UK.

One sentence summary: Development of a Golden Gate-based assembly standard for cloning and transformation in cyanobacteria that is compatible with, and builds on, the broadly established plant modular cloning syntax.

Author Contributions: RV, GARG, AAS and AM designed the study with input from BW; RV, GARG, AAS, AP, JM, MDG, KV performed experiments; VZ designed online software; BW contributed research reagents and materials. AM prepared the manuscript with input from experimentalists and BW, CJH and DL-S.

ABSTRACT (250 word limit)

Recent advances in synthetic biology research have been underpinned by an exponential increase in available genomic information and a proliferation of advanced DNA assembly tools. The adoption of plasmid vector assembly standards and parts libraries has greatly enhanced the reproducibility of research and exchange of parts between different labs and biological systems. However, a standardised Modular Cloning (MoClo) system is not yet available for cyanobacteria, which lag behind other prokaryotes in synthetic biology despite their huge potential in biotechnological applications. By building on the assembly library and syntax of the Plant Golden Gate MoClo kit, we have developed a versatile system called CyanoGate that unites cyanobacteria with plant and algal systems. Here we have generated a suite of parts and acceptor vectors for making i) marked/unmarked knock-outs or integrations using an integrative acceptor vector, and ii) transient multigene expression and repression systems using known and novel replicative vectors. We have tested and compared the CyanoGate system in the established model cyanobacterium *Synechocystis* sp. PCC 6803 and the more recently described fast-growing strain *Synechococcus elongatus* UTEX 2973. We observed that fast-growth phenotype in UTEX 2973 is only evident under specific growth conditions, but that UTEX 2973 can accumulate high levels of proteins with strong native or synthetic promoters. The system is publicly available and can be readily expanded to accommodate other standardised MoClo parts to accelerate the development of reliable synthetic biology tools for the cyanobacterial community.

INTRODUCTION

Much work is focused on expanding synthetic biology approaches to engineer photosynthetic organisms, including cyanobacteria. Cyanobacteria are an evolutionarily ancient and diverse phylum of photosynthetic prokaryotic organisms that are ecologically important, and are thought to contribute *ca*. 25% to oceanic net primary productivity (Castenholz, 2001; Flombaum et al., 2013). The chloroplasts of all photosynthetic eukaryotes, including plants, resulted from the endosymbiotic uptake of a cyanobacterium by a eukaryotic ancestor (Keeling, 2004). Therefore, cyanobacteria have proved useful as model organisms for the study of photosynthesis, electron transport and associated biochemical pathways, many of which are conserved in eukaryotic algae and higher plants. Several unique aspects of cyanobacterial photosynthesis, such as the biophysical carbon concentrating mechanism, also show promise as a means for enhancing productivity in crop plants (Rae et al., 2017). Furthermore, cyanobacteria are increasingly recognized as valuable platforms for industrial biotechnology

to convert CO_2 and H_2O into valuable products using solar energy (Tan et al., 2011; Ducat et al., 2011; Ramey et al., 2015). They are metabolically diverse and encode many components (e.g. P450 cytochromes) necessary for generating high-value pharmaceutical products that can be challenging to produce in other systems (Nielsen et al., 2016; Wlodarczyk et al., 2016; Pye et al., 2017; Stensjö et al., 2017). Furthermore, cyanobacteria show significant promise in biophotovoltaic devices for generating electrical energy (McCormick et al., 2015; Saar et al., 2018).

Based on morphological complexity, cyanobacteria are classified into five sub-sections (I-V) (Castenholz, 2001). Several members of the five sub-sections have been reportedly transformed (Vioque, 2007; Stucken et al., 2012), suggesting that many cyanobacterial species are amenable to genetic manipulation. Exogenous DNA can be integrated into or removed from the genome through homologous recombination-based approaches using natural transformation, conjugation (tri-parental mating), or electroporation (Heidorn et al., 2011). Exogenous DNA can also be propagated by replicative vectors, although the latter are currently restricted to the use of a single vector-type based on the broad-host range RSF1010 origin (Mermet-Bouvier et al., 1993; Huang et al., 2010; Taton et al., 2014). Transformation tools have been developed for generating "unmarked" mutant strains (lacking an antibiotic resistance marker cassette) in several model species, such as *Synechocystis* sp. PCC 6803 (*Synechocystis* hereafter) (Lea-Smith et al., 2016). More recently, markerless genome editing using CRISPR-based approaches has been demonstrated to function in both unicellular and filamentous strains (Ungerer and Pakrasi, 2016; Wendt et al., 2016).

Although exciting progress is being made in developing effective transformation systems, cyanobacteria still lag behind in the field of synthetic biology compared to bacterial (heterotrophic), yeast and mammalian systems. Relatively few broad-host range genetic parts have been characterised, but many libraries of parts for constructing regulatory modules and circuits are starting to become available, albeit using different standards, which makes them difficult to combine (Huang and Lindblad, 2013; Camsund et al., 2014; Albers et al., 2015; Markley et al., 2015; Englund et al., 2016; Kim et al., 2017; Immethun et al., 2017; Taton et al., 2017; Ferreira et al., 2018; Li et al., 2018; Liu and Pakrasi, 2018; Wang et al., 2018). One key challenge is clear: parts that are widely used in *Escherichia coli* behave very differently in model cyanobacterial species, such as *Synechocystis* (Heidorn et al., 2011). Furthermore, different cyanobacterial strains generally show a wide variation for functionality and

performance of different genetic parts (e.g. promoters, reporter genes and antibiotic resistance markers) (Taton et al., 2014; Englund et al., 2016; Kim et al., 2017; Taton et al., 2017). This suggests that parts need to be validated, calibrated, and perhaps modified, for individual strains, including model species and strains that may be more commercially relevant. Rapid cloning and assembly methods are essential for accelerating the 'design, build, test and learn' cycle, which is a central tenet of synthetic biology (Nielsen and Keasling, 2016).

The adoption of new cloning and vector assembly methods (e.g. Isothermal (Gibson) Assembly and MoClo), assembly standards and part libraries has greatly enhanced the scalability of synthetic biology-based approaches in a range of biological systems (Moore et al., 2016; Vazquez-Vilar et al., 2018). Recent advances in synthetic biology have led to the development of standards for Type IIS restriction endonuclease-mediated assembly (commonly known as Golden Gate cloning) for several model systems, including plants (Sarrion-Perdigones et al., 2013; Engler et al., 2014; Andreou and Nakayama, 2018). Based on a common Golden Gate Modular Cloning (MoClo) syntax, large libraries are now available for fusion of different genetic parts to assemble complex vectors cheaply and easily without proprietary tools and reagents (Patron et al., 2015). High-throughput and automated assembly are projected to be widely available soon through DNA synthesis and construction facilities, such as the UK DNA Synthesis Foundries, where MoClo is seen as the most suitable assembly standard (Chambers et al., 2016).

Here we have developed an easy-to-use system called CyanoGate that unites cyanobacteria with plant and algal systems, by building on the established Golden Gate MoClo syntax and assembly library for plants (Engler et al., 2014) that has been adopted by the OpenPlant consortium (www.openplant.org), iGEM competitions as "Phytobricks" and the MoClo kit for the microalga *Chlamydomonas reinhardtii* (Crozet et al., 2018). Firstly, we have constructed and characterised a suite of known and novel genetic parts (level 0) for use in cyanobacterial research, including promoters, terminators, antibiotic resistant markers, neutral sites and gene repression systems (Na et al., 2013; Yao et al., 2015; Sun et al., 2018). Secondly, we have designed an additional level of acceptor vectors (level T) to facilitate integrative or replicative transformation. We have characterised assembled level T vectors in *Synechocystis* and in *Synechococcus elongatus* UTEX 2973 (UTEX 2973 hereafter), which has a reported doubling time similar to that of *Saccharomyces cerevisiae* under specific growth conditions (Yu et al.,

2015; Ungerer et al., 2018a; 2018b). Lastly, we have developed an online tool for assembly of CyanoGate and Plant MoClo vectors to assist with the adoption of the CyanoGate system.

MATERIALS AND METHODS

Cyanobacterial culture conditions

Cyanobacterial strains of *Synechocystis*, UTEX 2973 and *Synechococcus elongatus* PCC 7942 (PCC 7942 hereafter) were maintained on 1.5% (w/v) agar plates containing BG11 medium. Liquid cultures were grown in Erlenmeyer flasks (100 ml) containing BG11 medium (Rippka et al., 1979) supplemented with 10 mM $NaHCO_3$, shaken at 100 rpm and aerated with filter-sterilised water-saturated atmospheric air. *Synechocystis* and PCC 7942 strains were grown at 30°C with continuous light (100 µmol photons m^{-2} s^{-1}) and UTEX 2793 strains were grown at 40°C with 300 µmol photons m^{-2} s^{-1} in an Infors Multitron-Pro supplied with warm white LED lighting (Infors HT).

Growth analysis

Growth of *Synechocystis*, UTEX 2973 and PCC 7942 was measured in a Photon Systems Instrument Multicultivator MC 1000-OD (MC). Starter cultures were grown in a Photon Systems Instrument AlgaeTron AG 230 at 30°C under continuous warm-white light (100 µmol photons m^{-2} s^{-1}) with air bubbling and shaken at 160 rpm unless otherwise indicated. These were grown to an optical density at 750 nm (OD_{750}) of approximately 1.0, and used to seed 80 ml cultures for growth in the MC at a starting OD_{720} of ~0.2 (the MC measures culture growth at OD_{720}). Cultures were then grown under continuous warm-white light at 30°C (300 µmol photons m^{-2} s^{-1}) with air bubbling or 30°C (under 300 or 500 µmol photons m^{-2} s^{-1}) or 41, 45 and 50°C (500 µmol photons m^{-2} s^{-1}) with 5% CO_2 bubbling until the fastest grown strain was at OD_{720} = ~0.9, the maximum accurate OD that can be measured with this device. A total of 5-6 replicate experiments were performed over two separate runs (16 in total), each inoculated from a different starter culture.

Vector construction

Level 0 vectors

Native cyanobacterial genetic parts were amplified from genomic DNA using NEB Q5 High-Fidelity DNA Polymerase (New England Biolabs) (**Fig. 1**; **Supplementary Table S1**). Where necessary, native genetic parts were domesticated (i.e. *Bsa*I and *Bpi*I sites were removed) using

specific primers. Alternatively, parts were synthesised as Gblocks® DNA fragments (Integrated DNA Technology) and cloned directly into an appropriate level 0 acceptor (see **Supplementary Information S1** for vector maps) (Engler et al. 2014).

Golden Gate assembly reactions were performed with restriction enzymes *Bsa*I (New England Biolabs) or *Bpi*I (Thermofisher), and T4 DNA ligase (Thermofisher) (see **Supplementary Information S2-S4** for detailed protocols). Vectors were transformed into One Shot TOP10 chemically competent *Escherichia coli* (Thermofisher) as per the manufacturer's instructions. Transformed cultures were grown at 37°C on [1.5% (w/v)] LB agar plates or in liquid LB medium shaking at 260 rpm, with appropriate antibiotic selection for level 0, 1, M and P vectors as outlined in Engler et al. (2014).

Level T acceptor vectors and new level 0 acceptors

A new level T vector system was designed that provides MoClo-compatible replicative vectors or integrative vectors for genomic modifications in cyanobacteria (Heidorn et al., 2011) (**Fig. 2; Supplementary Table S1; Supplementary Information S1**). For replicative vectors, we modified the pPMQAK1 carrying an RSF1010 replicative origin (Huang et al., 2010) to make pPMQAK1-T, and vector pSEVA421 from the Standard European Vector Architecture (SEVA) 2.0 database (seva.cnb.csic.es) carrying the RK2 replicative origin to make pSEVA421-T (Silva-Rocha et al., 2013). Replicative vector backbones were domesticated to remove native *Bsa*I and *Bpi*I sites where appropriate. The region between the BioBrick's prefix and suffix was then replaced by a *lacZ* expression cassette flanked by two *Bpi*I sites that produce overhangs TGCC and GGGA, which are compatible with the plant Golden Gate MoClo assembly syntax for level 2 acceptors (e.g. pAGM4673) (Engler et al., 2014). For integrative vectors, we domesticated a pUC19 vector backbone and introduced two *Bpi*I sites compatible with a level 2 acceptor (as above) to make pUC19A-T and pUC19S-T. In addition, we made a new low copy level 0 acceptor (pSC101 origin of replication) for promoter parts based on the BioBrick standard vector pSB4K5 (Liu et al., 2018). DNA was amplified using NEB Q5 High-Fidelity DNA Polymerase (New England Biolabs). All vectors were sequenced following assembly to confirm domestication and the integrity of the MoClo cloning site.

Level 0 parts for CRISPRi and srRNA

A nuclease deficient Cas9 gene sequence sourced from Addgene (www.addgene.org/44249/) was domesticated and assembled as a level 0 CDS part (**Supplementary Table S1; S2**) (Qi et

al., 2013). Five promoters of different strengths were truncated to the transcriptional start site (TSS) and cloned into a new level 0 acceptor vector with the unique overhangs GGAG and TAGC (**Fig. 1**). Two new level 0 parts with the unique overhangs GTTT and CGCT were generated for the sgRNA scaffold and srRNA HFQ handle (based on MicC) (Na et al., 2013), respectively. Assembly of level 1 expression cassettes proceeded by combining appropriate level 0 parts with a PCR product for either a srRNA or sgRNA (**Fig. 1**).

Cyanobacterial transformation and conjugation

Transformation with integrative level T vectors was performed as in Lea-Smith et al. (2016). For transformation by electroporation, cultures were harvested during the 'exponential' growth phase (OD_{750} of ~0.6) by centrifugation at 4,000 *g* for 10 min. The cell pellet was washed 3 times with 2 ml of sterile 1 mM HEPES buffer (pH 7.5), re-suspended in water with 3-5 µg of level T vector DNA and transferred into a 0.1 cm electroporation cuvette (Scientific Laboratory Suppliers). Re-suspended cells were electroporated using an Eppendorf 2510 electroporator (Eppendorf) set to 1200 V. Sterile BG-11 (1 ml) was immediately added to the electroporated cells. Following a 1 hr incubation at RT, the cells were plated on 1.5% (w/v) agar plates containing BG-11 with antibiotics at standard working concentrations to select for transformed colonies. The plates were sealed with parafilm and placed under 15 µmol photons m^{-2} s^{-1} light at 30°C for 1 day. The plates were then moved to 30 µmol photons m^{-2} s^{-1} light until colonies appeared. After 15-20 days, putative transformants were recovered and streaked onto new plates with appropriate antibiotics for further study.

Genetic modification by conjugation in *Synechocystis* was facilitated by an *E. coli* strain (HB101) carrying both mobilizer and helper vectors pRK2013 (ATCC® 37159™) and pRL528 (www.addgene.org/58495/), respectively (Tsinoremas et al., 1994). For UTEX 2973, conjugation was facilitated by a MC1061 strain carrying mobilizer and helper vectors pRK24 (www.addgene.org/51950/) and pRL528, respectively. Cultures of HB101 and OneShot TOP10 *E. coli* strains carrying level T cargo vectors were grown for approximately 15 hr with appropriate antibiotics. Cyanobacterial strains were grown to an OD_{750} of ~1. All bacterial cultures were washed three times with either fresh LB medium for *E. coli* or BG11 for cyanobacteria prior to use. *Synechocystis* cultures (100 µl, OD_{750} of 0.5-0.8) were conjugated by combining appropriate HB101 and the cargo strains (100 µl each) and plating onto HATF 0.45 µm transfer membranes (Merck Millipore) placed on LB: BG11 (1: 19) agar plates. For UTEX 2973 conjugations, appropriate MC1061 and the cargo strains (100 µl each) were

initially combined and incubated at 30ºC for 30 min, then mixed with UTEX 2973 cultures (100 µl, OD_{750} of 0.5-0.8) and incubated at 30ºC for 2 hr, and then plated onto transfer membranes as above. *Synechocystis* and UTEX 2973 transconjugates were grown under culturing conditions outlined above. Following growth on non-selective media for 24 hr, the membranes were transferred to BG11 agar plates supplemented with appropriate antibiotics. Colonies were observed within a week for both strains. Chlorophyll content of wild-type (WT) and mutant strains was calculated as in Lea-Smith et al. (2013).

Fluorescence assays

Transgenic strains maintained on agar plates containing appropriate antibiotics were used to inoculate 10 ml seed cultures that were grown to an optical density at 750 nm (OD_{750}) of approximately 1.0, as measured with a WPA Biowave II spectrometer (Biochrom). Seed cultures were diluted to an OD_{750} of 0.2, and 2 ml starting cultures were transferred to 24-well plates (Costar® Corning Incorporated) for experiments. *Synechocystis* and UTEX 2973 strains were grown in an Infors Multitron-Pro in the same culturing conditions described above. OD_{750} was measured using a FLUOstar OMEGA microplate reader (BMG Labtech). Fluorescence of eYFP for individual cells (10,000 cells per culture) was measured by flow cytometry using an Attune NxT Flow Cytometer (Thermofisher). Cells were gated using forward and side scatter, and median eYFP fluorescence was calculated from excitation/emission wavelengths 488 nm/515–545 nm (Kelly et al., 2018) and reported at 48 hr unless otherwise stated.

Cell counts, soluble protein and eYFP quantification

Synechocystis and UTEX 2973 strains were cultured for 48 hr as described above, counted using a haemocytometer and then harvested for soluble protein extraction. Cells were pelleted by centrifugation at 4,000 *g* for 15 min, re-suspended in lysis buffer [0.1 M potassium phosphate buffer (pH 7.0), 2 mM DTT and one Roche cOmplete EDTA-free protease inhibitor tablet per 10 ml (Roche Diagnostics)] and lysed with 0.5 mm glass beads (Thistle Scientific) in a TissueLyser II (Qiagen). The cell lysate was centrifuged at 18,000 g for 30 min and the supernatant assayed for soluble protein content using Pierce 660nm Protein Assay Reagent against BSA standards (Thermo Fisher Scientific). Extracts were subjected to sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) in a 4–12% (w/v) polyacrylamide gel (Bolt® Bis-Tris Plus Gel; Thermo Fisher Scientific) alongside a SeeBlue Plus2 Prestained protein ladder (Thermo), transferred to polyvinylidene fluoride (PVDF) membrane then probed with monoclonal anti-GFP serum (AbCAM) at 1: 1,000 dilution, followed by LI-COR IRDye

®800CW goat anti-rabbit IgG (LI-COR Inc.) at 1: 10,000 dilution, then viewed on an LI-COR Odyssey CLx Imager. eYFP protein content was estimated by immunoblotting using densitometry using LI-COR Lite Studio software v5.2 . Relative eYFP protein abundance was estimated by densitometry using LI-COR Lite Studio software v5.2.

Plasmid vector and genome copy number determination

The genome copy number and copy number of heterologous self-replicating plasmid vectors in *Synechocystis* was estimated using a real-time quantitative PCR (RT-qPCR) approach adapted from Zerulla et al. (2016). Cytoplasmic extracts containing total cellular DNA were harvested from *Synechocystis* cultures after 48 hr growth (OD_{750} = *ca*. 5) according to Zerulla et al. (2016). Cells in 10 ml of culture were pelleted by centrifugation at 4,000 *g* for 15 min, disrupted by shaking at 30 Hz for 10 min in a TissueLyser II with a mixture of 0.2 mm and 0.5 mm acid washed glass beads (0.35 g each), and then resuspended in dH_2O. The culture cell count was determined prior to harvest using a haemocytometer and checked again after cell disruption to calculate the efficiency of cell disruption. A standard curve based on a dilution series of vector DNA was generated and used for RT-qPCR analysis in parallel with extracts carrying the same vector. Two DNA fragments (*ca*. 1 kb) targeting two separate loci (*petB* and *secA*) were amplified from isolated genomic DNA from *Synechocystis* using standard PCR (Pinto et al., 2012). DNA mass concentrations were determined photometrically and the concentrations of DNA molecules was calculated from the known molecular mass. As above, a standard curve based on a dilution series of the two fragments was generated to estimate genome copy number in the extracts (Zerulla et al., 2016). The Ct of the extracts were then plotted against the linear portion of the standard curves to estimate plasmid vector copy number and genome copy number per cell. Oligonucleotides used are summarised in **Supplementary Table S3**.

Confocal laser scanning microscopy

Cultures were imaged using Leica TCS SP8 confocal microscopy (Leica Microsystems) with a water immersion objective lens (HCX APO L 20x/0.50 W). Excitation/emission wavelengths were 514 nm/527–546 nm for eYFP and 514 nm/605–710 nm for chlorophyll autofluorescence.

RESULTS AND DISCUSSION

Construction of the CyanoGate system

The CyanoGate system integrates with the two-part Golden Gate MoClo Plant Tool Kit, which can be acquired from Addgene [standardised parts (Kit #1000000047) and backbone acceptor vectors (Kit # 1000000044), (www.addgene.org)] (Engler et al., 2014). A comparison of the benefits of MoClo- and Gibson assembly-based cloning strategies is shown in **Supplementary Information S5.** The syntax for level 0 parts was adapted for prokaryotic cyanobacteria to address typical cloning requirements for cyanobacterial research (**Fig. 1**). New level 0 parts were assembled from a variety of sources (**Supplementary Table S1**). Level 1, M and P acceptor vectors were adopted from the MoClo Plant Tool Kit, which facilitates assembly of level 0 parts in a level 1 vector, and subsequently up to seven level 1 modules in level M. Level M assemblies can be combined further into level P and cycled back into level M to produce larger multi-module vectors if required (**Supplementary Information S3**). Vectors >50 kb in size assembled by MoClo have been reported (Werner et al., 2012). Modules from level 1 or level P can be assembled in new level T vectors designed for cyanobacterial transformation (**Fig. 2**). We found that both UTEX 2973 and *Synechocystis* produced recombinants following electroporation or conjugation methods with level T vectors. For the majority of the work outlined below, we relied on the conjugation approach.

Integration - generating marked and unmarked knockout mutants

A common method for engineering stable genomic knock out and knock in mutants in several cyanobacteria relies on homologous recombination via integrative (suicide) vectors using a two-step marked-unmarked strategy (Lea-Smith et al., 2016) (**Supplementary Information S4**). Saar et al., 2018 recently used this approach to introduce up to five genomic alterations into a single *Synechocystis* strain. Firstly, marked mutants are generated with an integrative vector carrying two sequences (approximately 1 kb each) identical to the regions of the cyanobacterial chromosome flanking the deletion/insertion site. Two gene cassettes are inserted between these flanking sequences: a levansucrase expression cassette (*sacB*) that confers sensitivity to transgenic colonies grown on sucrose and an antibiotic resistance cassette (Ab^R) of choice. Secondly, unmarked mutants (carrying no selection markers) are generated from fully segregated marked lines using a separate integrative vector carrying only the flanking sequences and selection on plates containing sucrose.

We adapted this approach for the CyanoGate system (**Fig. 1**). To generate level 1 vectors for making knock out mutants, sequences flanking the upstream (UP FLANK) and downstream (DOWN FLANK) site of recombination were ligated into the plant MoClo Prom+5U (with

overhangs GGAG-AATG), and 3U+Ter (GCTT-CGCT) positions, respectively, to generate new level 0 parts (**Fig. 1B**). In addition, full expression cassettes were made for sucrose selection (*sacB*) and antibiotic resistance (Ab^RSpec, Ab^RKan and Ab^REry) in level 0 that ligate into positions SP (AATG-AGGT) and CDS2 (stop) (AGGT-GCTT), respectively. Marked level 1 modules can be assembled using UP FLANK, DOWN FLANK, *sacB* and the required Ab^R level 0 part. For generating the corresponding unmarked level 1 module, a short 59 bp linker (UNMARK LINKER) can be ligated into the CDS1ns (AATG-GCTT) position for assembly with an UP FLANK and DOWN FLANK (**Fig. 1D**). Unmarked and marked level 1 modules can then be assembled into level T integrative vectors, with the potential capacity to include multiple knockout modules in a single level T vector.

To validate our approach, we constructed the level 0 flanking vectors pC0.024 and pC0.025 and assembled two level T integrative vectors using pUC19-T (*cpcBA*-M and *cpcBA*-UM, with and without the *sacB* and Ab^R cassettes, respectively) to remove the *cpcBA* promoter and operon in *Synechocystis* and generate an "Olive" mutant unable to produce the phycobiliprotein C-phycocyanin (Kirst et al., 2014; Lea-Smith et al., 2014) (**Fig. 3**; **Supplementary Table S1**). Following transformation with *cpcBA*-M, we successfully generated a marked Δ*cpcBA* mutant carrying the *sacB* and the Ab^RKan cassettes after selective segregation (*ca*. 3 months) (**Fig. 3A**). The unmarked Δ*cpcBA* mutant was then isolated following transformation of the marked Δ*cpcBA* mutant with *cpcBA*-UM and selection on sucrose (*ca*. 2 weeks) (**Fig. 3B**). Absence of C-phycocyanin in the Olive mutant resulted in a characteristic drop in absorbance at 625 nm (**Fig. 3D**) and a significant reduction in chlorophyll content compared to WT cells (28.4 ± 0.2 and 48.3 ± 0.2 amol chl $cell^{-1}$, respectively) (Kirst et al., 2014; Lea- Smith et al., 2014).

Generating knock in mutants

Flexibility in designing level 1 insertion cassettes is needed when making knock in mutants. Thus, for knock in mutants the upstream and downstream sequences flanking the insertion site, and any required expression or marker cassettes, are first assembled into separate level 1 modules from UP FLANK and DOWN FLANK level 0 parts (**Fig. 1E, F**). Seven level 1 modules can be assembled directly into Level T (**Fig. 2**), thus with a single pair of flanking sequences up to five level 1 expression cassettes could be included in a Level T vector.

Linker parts (20 bp) UP FLANK LINKER and DOWN FLANK LINKER were generated to allow assembly of level 0 UP FLANK and DOWN FLANK parts into separate level 1 acceptor

vectors. Similarly, level 0 linker parts were generated for *sacB* and Ab^R (**Fig. 1H, I**). Level 1 vectors at different positions can then be assembled in level T (or M) containing one or more expression cassettes, an Ab^R of choice, or both *sacB* and Ab^R (**Fig. 2**).

Using this approach, CyanoGate can facilitate the generation of knock in mutants using a variety of strategies. For example, if retention of a resistance marker is not an experimental requirement (e.g. Liberton et al., 2017), only a single antibiotic resistance cassette needs to be included in level T. Alternatively, a two-step marked-unmarked strategy could be followed, as for generating knockout mutants.

While knock out strategies can target particular loci, knock in approaches often rely on recombination at designated 'neutral sites' within the genome of interest that can be disrupted with no or minimal impact on growth phenotype (Ng et al., 2015; Pinto et al., 2015). Based on loci reported in the literature, we have assembled a suite of flanking regions to target neutral sites in *Synechocystis* (designated 6803 NS1-4) (Pinto et al., 2015), *Synechococcus* sp. PCC 7002 (PCC 7002 hereafter) (designated 7002 NS1 and NS2) (Ruffing et al., 2016; Vogel et al., 2017), and neutral sites common to UTEX 2973, PCC 7942 and *Synechococcus elongatus* PCC 6301 (designated 7942 NS1-3) (Bustos and Golden, 1992; Kulkarni and Golden, 1997; Andersson et al., 2000; Niederholtmeyer et al., 2010) (**Supplementary Table S1**). Pinto et al. (2015) have qualitatively compared the impact of the four *Synechocystis* neutral sites assembled here under several different growth conditions, and observed that insertions at 6803 NS3 and NS4 had no significant effect on growth compared to WT cultures, while insertions at NS2 and NS1 had small but significant effects depending on the growth conditions. Several studies have used 6803 NS3, for example, to engineer a *Synechocystis* strain for the bioremediation of microcystins (Dexter et al., 2018) and the development of T7 polymerase-based synthetic promoter systems (Ferreira et al., 2018). For the two PCC 7002 neutral sites, growth rates with insertions at 7002 NS1 were slightly reduced (Vogel et al., 2017), but not significantly affected with insertions at 7002 NS2 (Ruffing et al., 2016). Insertions at the three 7942 neutral sites reportedly have no phenotypic effect on morphology or growth rate (Clerico et al., 2007; Niederholtmeyer et al., 2010) and have been used to study mRNA stability and translation (Kulkarni and Golden, 1997), circadian rhythms (Anderson et al. 2000), chromosome duplication (Watanabe et al., 2017) and to engineer PCC 7942 for synthesising heterologous products (Niederholtmeyer et al., 2010; Gao et al., 2016). When using the neutral sites supplied with CyanoGate (or others), we would still recommend a thorough growth

analysis under the specific culturing conditions being tested to identify any potential impact of the inserted DNA on growth phenotype.

To validate our system, we generated a level T vector carrying the flanking regions for the *cpcBA* operon and an eYFP expression cassette (*cpcBA*-eYFP) (**Fig. 4A, B; Supplementary Table S1**). We successfully transformed this vector into our marked "Olive" *Synechocystis* mutant, to generate a stable olive mutant with constitutive expression of eYFP (Olive-eYFP) (**Fig. 4C**).

Expression comparison for promoter parts in *Synechocystis* and UTEX 2973

We constructed level 0 parts for a wide selection of synthetic promoters and promoters native to *Synechocystis*. Promoters were assembled as expression cassettes driving eYFP in replicative level T vector pPMQAK1-T to test for differences in expression when conjugated into *Synechocystis* or UTEX 2973. We first compared the growth rates of *Synechocystis*, UTEX and PCC 7942 [a close relative of UTEX 2973 (Yu et al., 2015)] under a variety of different culturing conditions (**Supplementary Figure S1**). We found that growth rates were comparable between *Synechocystis* and PCC 7942 at temperatures below 40ºC regardless of light levels and supplementation with CO_2. In contrast, UTEX 2973 grew poorly under those conditions. UTEX 2973 only showed an enhanced growth rate at 45ºC under the highest light tested (500 µmol photons m^{-2} s^{-1}) with CO_2, while all three strains failed to grow at 50ºC. These results confirm that the enhanced growth phenotype reported for UTEX 2973 requires specific conditions as reported by Ungerer et al. (2018a; 2018b). Furthermore, they are consistent with recent reports that this phenotype is linked to an increased stress tolerance, which has been attributed to a small number of nucleotide polymorphisms (Lou et al., 2018; Ungerer et al., 2018b). We proceeded with CyanoGate part characterisations and comparisons under the best conditions achievable for *Synechocystis* and UTEX 2973 (see Materials and Methods) (**Supplementary Figure S2A**).

Promoters native to Synechocystis

We assembled several previously reported promotors from *Synechocystis* in the CyanoGate kit. These include six inducible/repressible promoters (P_{nrsB}, $P_{coaT,}$ P_{nirA}, P_{petE}, P_{isiAB}, and P_{arsB}), which were placed in front of the strong, synthetic *Synechocystis* ribosomal binding site (RBS^{*}) (Heidorn et al., 2011) as used in Englund et al. (2016) (**Supplementary Table S1; Supplementary Information S1**). P_{nrsB} and P_{coaT} drive the expression of nickel and cobalt ion

efflux pumps and are induced by Ni^{2+}, and Co^{2+} or Zn^{2+} , respectively (Peca et al., 2008; Blasi et al., 2012; Guerrero et al., 2012; Englund et al., 2016). P_{nirA}, from the nitrate assimilation operon, is induced by the presence of NO_3^- and/or NO_2^- (Kikuchi et al., 1996; Qi et al., 2005). P_{petE} drives the expression of plastocyanin and is induced by Cu^{2+}, which has previously been used for the expression of heterologous genes (Guerrero et al., 2012; Camsund et al., 2014). The promoter of the *isiAB* operon (P_{isiAB}) is repressed by Fe^{3+} and activated when the cell is under iron stress (Kunert et al., 2003). P_{arsB} drives the expression of a putative arsenite and antimonite carrier and is activated by AsO_2^- (Blasi et al., 2012).

We also cloned the *rnpB* promoter, P_{rnpB}, from the Ribonuclease P gene (Huang et al., 2010), a long version of the *psbA2* promoter, P_{psbA2L}, from the Photosystem II protein D1 gene (Lindberg et al., 2010; Englund et al., 2016) and the promoter of the C-phycocyanin operon, P_{cpc560} (also known as P_{cpcB} and P_{cpcBA}) (Zhou et al., 2014). P_{rnpB} and P_{psbA2L} were placed in front of RBS* (Heidorn et al., 2011) (**Fig. 5A**). To build on a previous functional characterisation of P_{cpc560} (Zhou et al., 2010), we assembled four variants of this strong promoter. Firstly, $P_{cpc560+A}$ consisted of the promoter and the 4 bp MoClo overhang AATG. Secondly, P_{cpc560} was truncated by one bp (A), so that that the start codon was aligned with the native P_{cpc560} RBS spacer region length. Zhou et al. (2014) identified 14 predicted transcription factor binding sites (TFBSs) in the upstream region of P_{cpc560} (-556 to -381 bp) and removal of this region resulted in a significant loss of promoter activity. However, alignment of the reported TFBSs showed their locations are in the downstream region of the promoter (-180 to -5 bp). We identified 11 additional predicted TFBSs using Virtual Footprint (Munch et al., 2005) in the upstream region and hypothesised that the promoter activity may be modified by duplicating either of these regions. So thirdly, we generated P_{cpc560_Dx2} containing a duplicated downstream TFBS region. For P_{cpc560_Dx2}, only the region between -31 to -180 bp was duplicated to avoid repeating the Shine-Dalgarno (SD) sequence. Fourthly, we duplicated the upstream region to generate P_{cpc560_Ux2}. We then assembled P_{rnpB}, P_{psbA2L} and the four P_{cpc560} variants with eYFP and the *rrnB* terminator (T_{rrnB}) into a Level 1 expression cassette, and subsequently into a level T replicative vector (pPMQAK1-T) for expression analysis (**Supplementary Table S2**).

In *Synechocystis* the highest expressing promoter was P_{cpc560} (**Fig. 5B**), which indicated that maintaining the native RBS spacer region for P_{cpc560} is important for maximising expression. Neither P_{cpc560_Dx2} nor P_{cpc560_Ux2} had higher expression levels compared to P_{cpc560}. P_{cpc560_Dx2} was strongly decreased compared to P_{cpc560}, suggesting that promoter function is sensitive to

modification of the downstream region and this region could be a useful target for modulating P_{cpc560} efficacy. Previous work in *Synechocystis* has suggested that modification of the middle region of P_{cpc560} (-380 to -181 bp) may also affect function (Lea-Smith et al., 2014). P_{psbA2L} produced lower expression levels than any variant of P_{cpc560} in *Synechocystis*, while P_{rnpB} produced the lowest expression levels. The observed differences in expression levels are consistent with those in other studies with *Synechocystis* (Camsund et al., 2014; Englund et al., 2016; Liu and Pakrasi, 2018).

In UTEX 2973, the trend in expression patterns was similar to that in *Synechocystis* (**Fig. 5B**). However, the overall expression levels of eYFP measured in UTEX 2973 were significantly higher than in *Synechocystis*. P_{cpc560} was increased by 30%, while P_{rnpB} showed a 20-fold increase in expression relative to *Synechocystis*. The relative expression strength of P_{psbA2L} was also higher than in *Synechocystis*, and second only to P_{cpc560} in UTEX 2973. As promoters derived from P_{psbA} are responsive to increasing light levels (Englund et al., 2016), the increased levels of expression for P_{psbA2L} may be associated with the higher light intensities used for growing UTEX 2973 compared to *Synechocystis*. Background fluorescence levels were similar between UTEX 2973 and *Synechocystis* conjugated with an empty pPMQAK1-T vector (i.e. lacking an eYFP expression cassette), which suggested that the higher fluorescence values in UTEX 2973 were a direct result of increased levels of eYFP protein.

Heterologous and synthetic promoters

A suite of twenty constitutive synthetic promoters was assembled in level 0 based on the modified BioBricks BBa_J23119 library of promoters (Markley et al., 2015), and the synthetic P_{trc10}, P_{tic10} and P_{tac10} promoters (Huang et al., 2010; Albers et al., 2015) (**Supplementary Table S1; Supplementary Information S1**). We retained the broad-range BBa_B0034 RBS (AAAGAGGAGAAA) and *lac* operator (*lacO*) from Huang et al. (2010), for future *lacI*-based repression experiments (*lacI* and the P_{lacIQ} promoter are included in the CyanoGate kit) (Bhal et al., 1977). We cloned eight new variants (J23119MH_V01-8) with mutations in the canonical BBa_J23119 promoter sequence (**Fig. 6A**). Additionally, we included the L-arabinose-inducible promoter from *E. coli* (P_{BAD}) (Abe et al., 2014).

We encountered an unexpected challenge with random internal deletions in the -35 and -10 regions of some promoters of the BBa_J23119 library and *trc* promoter variants when cloning them into level 0 acceptors. Similar issues were reported previously for the *E. coli* EcoFlex kit

(Moore et al., 2016) that may relate to the functionality of the promoters and the host vector copy number in *E. coli*, which consequently resulted in cell toxicity and selection for mutated promoter variants. To resolve this issue, we generated a low copy level 0 promoter acceptor vector compatible with CyanoGate (pSB4K5 acceptor) for cloning recalcitrant promoters (**Supplementary Table S1; Supplementary Information S1**). Subsequent assemblies in level 1 and T showed no indication of further mutation.

We then tested the expression levels of eYFP driven by the synthetic promoters in *Synechocystis* and UTEX 2973 following assembly in pPMQAK1-T (**Fig. 6B; Supplementary Table S2**). The synthetic promoters showed a 120-fold dynamic range in both cyanobacterial strains. Furthermore, a similar trend in promoter expression strength was observed ($R^2 = 0.84$) (**Fig. 6C)**. However, eYFP expression levels were on average eight-fold higher in UTEX 2973 compared to *Synechocystis*. In *Synechocystis*, the highest expression levels were observed for J23119 and P_{trc10}, but these were still approximately 50% lower than values for the native P_{cpc560} promoters (**Fig. 5B**). The expression trends for the BBa_J23119 library were consistent with the subset reported by Camsund et al. (2014) in *Synechocystis*, while the observed differences between P_{trc10} and P_{cpc560} were similar to those reported by Liu and Pakrasi (2018).

In contrast, the expression levels in UTEX 2973 for J23119 were approximately 50% higher than P_{cpc560}. Several synthetic promoters showed expression levels in a similar range to those for the native P_{cpc560} promoter variants, including P_{trc10}, J23111 and the J23119 variant V02. V02 is identical to J23111 except for an additional 'G' between the -35 and -10 motifs, suggesting that small changes in the length of this spacer region may not be critical for promoter strength (similar expression levels were also observed for these two promoters in *Synechocystis*). In contrast, a single bp difference between J23111 and J23106 in the -35 motif resulted in an eight- and ten-fold reduction in expression in *Synechocystis* and UTEX 2973, respectively. The results for UTEX 2973 were unexpected, and to our knowledge no studies to date have directly compared these promoters in this strain. Recent work has examined the expression of *β*-galactosidase using promoters such as P_{cpc560} and P_{trc} in UTEX 2973 (Li et al., 2018). Li et al. (2018) highlighted that different growth environments (e.g. light levels) can have significant effects on protein expression. Changes in culture density can also affect promoter activity, such that protein expression levels can change during the exponential and stationary growth stages depending on the promoter and expression vector used (Ng et al., 2015; Madsen et al., 2018). Here we tracked eYFP expression levels over time for three days

during early and late exponential growth phase for *Synechocystis* and UTEX 2973. Although expression levels for each promoter fluctuated over time, with peak expression levels at 24 hr and 48 hr in UTEX 2973 and *Synechocystis*, respectively, the overall expression trends were generally consistent for the two strains (**Supplementary Figure S2B**).

Protein expression levels in *Synechocystis* and UTEX 2973

To investigate further the increased levels of eYFP expression observed in UTEX 2973 compared to *Synechocystis*, we examined cell morphology, protein content and eYFP protein abundances in expression lines for each strain. Confocal image analysis confirmed the coccoid and rod shapes of *Synechocystis* and UTEX 2973, respectively, and the differences in cell size (van de Meene et al. 2006; Yu et al., 2015) (**Fig. 7A**). Immunoblot analyses of eYFP from protein extracts of four eYFP expressing strains correlated well with previous flow cytometry measurements (**Fig 5; 6**). eYFP driven by the J23119 promoter in UTEX 2973 produced the highest levels of eYFP protein (**Fig. 7B, C**). Although the density of cells in culture was two-fold higher in *Synechocystis* compared to UTEX 2973 (**Fig. 7D)**, the protein content per cell was six-fold lower (**Fig. 7E**). We then estimated the average cell volumes for *Synechocystis* and UTEX 2973 at 3.91 ± 0.106 μm^3 and 6.83 ± 0.166 μm^3 (n = 50 each), respectively, based on measurements from confocal microscopy images (**Supplementary Figure S3**). Based on those estimates, we calculated that the density of soluble protein per cell was four-fold higher in UTEX 2973 compared to *Synechocystis* (**Fig. 7F**). Thus, we hypothesised that the enhanced levels of eYFP observed in UTEX 2973 were a result of the expression system harnessing a larger available amino acid pool. Mueller et al. (2017) have reported that UTEX 2973 has an increased investment in amino acid content compared to PCC 7942, which may be linked to higher rates of translation in UTEX 2973. Therefore, UTEX 2973 continues to show promise as a bioplatform for generating heterologous protein products, although future work should study production rates under conditions optimal for faster-growth (Lou et al., 2018; Ungerer et al., 2018a). Recent characterisation of the UTEX 2973 transcriptome will also assist with native promoter characterisations (Tan et al., 2017).

The RK2 origin of replication is functional in *Synechocystis*

Synthetic biology tools (e.g. gene expression circuits, CRISPR/Cas-based systems) are often distributed between multiple plasmid vectors at different copy numbers in order to synthesise each component at the required concentration (Bradley et al., 2016). The large RSF1010 vector is able to replicate in a broad range of microbes including gram-negative bacteria such as *E.*

coli and several cyanobacterial species. However, for 25 years it has remained the only non-native vector reported to be able to self-replicate in cyanobacteria (Mermet-Bouvier et al., 1993). Recently, two small plasmids native to *Synechocystis*, pCA2.4 and pCB2.4, have been engineered for gene expression (Armshaw et al., 2015; Ng et al., 2015; Liu and Pakrasi, 2018). The pANS plasmid (native to PCC 7942) has also been adapted as a replicative vector, but so far it has been only shown to function in PCC 7942 and *Anabaena* PCC 7120 (Chen et al., 2016). Similarly, the high copy number plasmid pAQ1 (native to PCC 7002) has been engineered for heterologous expression, but up to now is has only been used in PCC 7002 (Xu et al., 2011). To expand the replication origins available for cyanobacterial research further we tested the capacity for vectors from the SEVA library to replicate in *Synechocystis* (Silva-Rocha et al., 2013).

We acquired three vectors driven by three different replication origins [pSEVA421 (RK2), pSEVA431 (pBBR1) and pSEVA442 (pRO1600/ColE1)] and carrying a spectinomycin antibiotic resistance marker. These vectors were domesticated and modified as level T acceptor vectors, assembled and then transformed into *Synechocystis* by electroporation or conjugation. Only *Synechocystis* strains conjugated with vectors carrying RK2 (pSEVA421-T) grew on spectinomycin-containing plates (**Supplementary Table S1; Supplementary Information S1**). To confirm that RSF1010 and RK2 replication origins can replicate autonomously in *Synechocystis*, we recovered the pPMQAK1-T or pSEVA421-T vector from lysates of axenic *Synechocystis* strains previously conjugated with each vector by transformation into *E. coli*. The identity and integrity of pPMQAK1-T and pSEVA421-T extracted from transformed *E. coli* colonies were confirmed by restriction digest and Sanger sequencing.

We then assembled two level T vectors with an eYFP expression cassette (P_{cpc560}-eYFP- T_{rrnB}) to produce pPMQAK1-T-eYFP and pSEVA421-T-eYFP, which were conjugated into *Synechocystis* (**Fig. 8; Supplementary Table S2**). Both pPMQAK1-T-eYFP and pSEVA421-T-eYFP transconjugates grew at similar rates in 50 µg ml^{-1} kanamycin and 5 µg ml^{-1} spectinomycin, respectively (**Fig. 8A**). However, eYFP levels were 8-fold lower in pSEVA421-T-eYFP, suggesting that RK2 has a reduced copy number relative to RSF1010 in *Synechocystis* (**Fig. 8B**). We measured the heterologous plasmid vector copy number in strains expressing pSEVA421-T or pPMQAK1-T and estimated an average copy number per cell of 9 ± 2 and 31 ± 5, respectively (**Fig. 8C**). The copy number for pPMQAK1-T was similar to values reported previously for RSF1010-derived vectors in *Synechocystis* (*ca*. 30) (Ng et al.,

2000). Our results are also consistent with the lower copy numbers in *E. coli* for vectors with RK2 (4-7 copies) compared to RSF1010 (10-12 copies) replication origins (Frey et al., 1992; Blasina et al., 1996). Furthermore, we compared the genome copies per cell between transformants and wild-type strains and found no significant differences - the average value was 11 ± 2, which is consistent with the typical range of genome copy numbers observed in *Synechocystis* cells (Zerulla et al., 2016).

Gene repression systems

CRISPR (clustered regularly interspaced short palindromic) interference (CRISPRi) is a relatively new but well characterised tool for modulating genes expression at the transcription stage in a sequence-specific manner (Qi et al., 2013; Behler et al., 2018). CRISPRi typically uses a nuclease deficient Cas9 from *Streptococcus pyogenes* (dCas9) and has been demonstrated to work in several cyanobacterial species, including *Synechocystis* (Yao et al., 2015), PCC 7002 (Gordon et al., 2016); PCC 7942 (Huang et al., 2016) and *Anabaena* sp. PCC 7120 (Higo et al., 2018). A second approach for gene repression uses rationally designed small regulatory RNAs (srRNAs) to regulate gene expression at the translation stage (Na et al., 2013; Higo et al., 2016). The synthetic srRNA is attached to a scaffold to recruit the Hfq protein, an RNA chaperone that is conserved in a wide-range of bacteria and cyanobacteria, which facilitates the hybridization of srRNA and target mRNA, and directs mRNA for degradation. The role of cyanobacterial Hfq in interacting with synthetic srRNAs is still unclear (Zess et al., 2016). However, regulatory ability can be improved by introducing Hfq from *E. coli* into *Synechocystis* (Sakai et al., 2015). Both CRISPRi- and srRNA-based systems have potential advantages as they can be used to repress multiple genes simultaneously.

To validate the CRISPRi system, we assembled an expression cassette for dCas9 (P_{cpc560}-dCas9-T_{rrnB}) on the Level 1 position 1 vector pICH47732, and four different sgRNA expression cassettes (P_{trc10_TSS}-sgRNA-sgRNA scaffold) targeting eYFP on the Level 1 position 2 vector pICH47742 (Engler et al., 2014) (**Supplementary Table S2**). For assembly of CRISPRi sgRNA expression cassettes in level 1, we targeted four 18-22 bp regions of the eYFP non-template strand with an adjacent 3' protospacer adjacent motif (PAM) of 5′-NGG-3′, as required by *S. pyogenes* dCas9 (**Fig. 9A**). The sgRNA sequences contained no off-target sites in the *Synechocystis* genome (confirmed by CasOT; Xiao et al., 2014). The sgRNAs were made by PCR using two complementary primers carrying the required overhangs and *Bsa*I sites, and were assembled with P_{trc10}_TSS promoter (pC0.203) and the sgRNA scaffold (pC0.122) (**Fig.**

1K). Level T vectors were assembled carrying dCas9 and a single sgRNA, or just the sgRNA alone. We subsequently conjugated the Olive-eYFP mutant and tracked eYFP expression.

Transconjugates carrying only the sgRNA showed no reduction in eYFP level compared to non-transconjugated Olive-eYFP (**Fig. 9B**). However, all strains carrying dCas9 and a sgRNA showed a decrease in eYFP that ranged from 40-90% depending on the sgRNA used. These reductions are similar to those observed previously in PCC 7002 and in *Synechocystis* (Yao et al., 2016; Gordon et al., 2016) and demonstrated that CRISPRi system is functional in the CyanoGate kit.

CONCLUSION

The CyanoGate kit was designed to increase the availability of well characterised libraries and standardised modular parts in cyanobacteria (Sun et al., 2018). We aimed to simplify and accelerate modular cloning methods in cyanobacterial research and allow integration with the growing number of labs that rely on the established common plant and algal syntax for multi-part DNA assembly (Patron et al., 2015; Crozet et al., 2018). Here, we have demonstrated the functionality of CyanoGate in sufficient detail to show that it is straightforward to adopt and functionally robust across two different cyanobacterial species. CyanoGate includes parts for usage in other cyanobacterial species and could likely be utilised also in non-cyanobacterial microbes amenable to transformation (e.g. *Rhodopseudomonas* spp.) and adapted for use in subcellular eukaryotic compartments of prokaryote origin (e.g. chloroplasts) (Economou et al., 2014; Doud et al., 2017; Leonard et al., 2018). In addition to the parts discussed, we have also assembled a suite of 21 terminators (**Supplementary Table S1**). To increase the accessibility and usability of the CyanoGate, we have included the vector maps for all parts and new acceptors (**Supplementary Information S1**), implemented support for Cyanogate assemblies in the online DNA "Design and Build" portal of the Edinburgh Genome Foundry (dab.genomefoundry.org) (**Supplementary Information S6**), and submitted all vectors as a toolkit for order from Addgene (Addgene Kit #1000000146; www.addgene.org/kits/mccormick-cyanogate).

Standardisation will help to accelerate the development of reliable synthetic biology tools for biotechnological applications and promote sharing and evaluation of genetic parts in different species and under different culturing conditions (Patron et al., 2015). Going forward, it will be important to test the performance of different parts with different components (e.g. gene

expression cassettes) and in different assembly combinations. Several groups using plant MoClo assembly have reported differences in cassette expression and functionality depending on position and orientations (e.g. Ordon et al., 2017), which highlights a key synthetic biology crux - the performance of a system is not simply the sum of its components (Mutalik et al., 2013; Heyduk et al., 2018).

The increasing availability of genome-scale metabolic models for different cyanobacterial species and their utilisation for guiding engineering strategies for producing heterologous high-value biochemicals has helped to re-invigorate interest in the industrial potential of cyanobacteria (Knoop et al., 2013; Hendry et al., 2016; Mohammadi et al., 2016; Shirai et al., 2016). Future efforts should focus on combining genome-scale metabolic models with synthetic biology approaches, which may help to overcome the production yield limitations observed for cyanobacterial cell factories (Nielsen et al., 2016), and will accelerate the development of more complex and precise gene control circuit systems that can better integrate with host metabolism and generate more robust strains (Bradley and Wang 2015; Jusiak et al., 2016; Luan and Lu, 2018). The future development of truly 'programmable' photosynthetic cells could provide significant advancements in addressing fundamental biological questions and tackling global challenges, including health and food security (Dobrin et al., 2016; Medford and Prasad, 2016; Smanksi et al., 2016).

SUPPLEMENTARY DATA

Supplementary Table S1. Table of all parts from CyanoGate kit generated in this work. Domestication refers to the removal of *BsaI* and/or *BsiI* sites (modifications are indicated in sequence maps provided in **Supp. Info. 2**). See separate .xlsx.

Supplementary Table S2. List of level T vectors used in this study.

Supplementary Table S3. Sequences of synthetic oligonucleotides used to determine copy number. Primers used for amplifying the *petB* locus were from Pinto et al. (2012).

Supplementary Information S1. Sequence maps (.gb files) of the components of the CyanoGate kit. See .zip file.

Supplementary Information S2. Protocols for MoClo assembly in level -1 through to level T. Protocols for assembly in level 0, level M and level T acceptor vectors (restriction enzyme *Bpi*I required, left). Protocols for assembly in level -1, level 1 and level P backbone vectors (restriction enzyme *Bsa*I required, right). Adapted from "A quick guide to Type IIS cloning" (Patron Lab; patronlab.org). For troubleshooting Type IIS mediated assembly we recommend synbio.tsl.ac.uk/docs.

Supplementary Information S3. Detailed assembly strategies using the CyanoGate kit.

Supplementary Information S4. Integrative engineering strategies using the CyanoGate kit. (A) Marked mutants are generated using a level T marked knock out vector carrying DNA sequences flanking the target locus of the chromosome (~1 kb), an antibiotic resistance cassette (Ab^R) and a sucrose selection cassette (*sacB*) that produces the toxic compound levansucrase in the presence of sucrose (20). Several rounds of segregation are required to identify a marked mutant. (B) Marked mutants then can be unmarked with a level T unmarked knock out vector and selection on sucrose-containing agar plates. (C) Unmarked knock in mutants can also be generated from marked mutants using a level T unmarked knock in vector carrying a gene expression cassette (UP FLANK LINKER and DOWN FLANK LINKER are shown in pink and light green, respectively). (D) Alternatively marked knock in mutants can be engineered in a single step using a level T marked knock vector (Ab^R UP LINKER and DOWN LINKER are shown in blue and orange, respectively). See **Fig. 2** for abbreviations.

Supplementary Information S5. Comparison of Gibson Assembly and Golden Gate Assembly. (A) A comparison of Gibson Assembly and Golden Gate Assembly pathways for building the level T vector cpcBA-eYFP described in **Fig. 4**. (B) Advantages and disadvantages of Gibson Assembly and Golden Gate Assembly.

Supplementary Information S6. Protocol and online interface for building CyanoGate vector assemblies. A CyanoGate online vector assembly tool called Design and Build (DAB) from the Edinburgh Genome Foundry.

Supplementary Figure S1. Comparison of growth for *Synechocystis*, PCC 7942 and UTEX 2973 under different culturing conditions. Values are the means ± SE from at least five biological replicates from two independent experiments.

Supplementary Figure S2. Growth and expression levels of heterologous and synthetic promoters in *Synechocystis* and UTEX 2973. (A) *Synechocystis* and UTEX 2973 was cultured for 72 hr at 30°C with continuous light (100 µmol photons m^{-2} s^{-1}) and 40°C with 300 µmol photons m^{-2} s^{-1}, respectively (see **Fig. 6**). Expression levels of eYFP are shown at three time points (24, 48 and 72 hr after inoculation). Values are the means ± SE from at least four biological replicates where each replicate represents the median measurements of 10,000 cells.

Supplementary Figure S3. Cell volume calculations for *Synechocystis* and UTEX 2973 from confocal microscopy images. (A) Example confocal images of *Synechocystis* (left) and UTEX 2973 (right) cells expressing eYFP driven by the J23119 promoter at 48 hr. Individual cells were selected and measured using Leica AF Lite software (Leica Microsystems). Top panel: eYFP fluorescence (green); middle panel: chlorophyll auto florescence (red); bottom panel: overlay of eYFP and chlorophyll signals (yellow). (B) Volume estimations based on confocal image data (n =50) (C) Mathematical formulas used for calculating cell volume based on the cell shapes of *Synechocystis* (coccus, spherical) and UTEX 2973 (bacillus, cylindrical).

ACKNOWLEDGEMENTS

We thank Conrad Mullineaux (Queen Mary University), Julie Zedler and Poul Erik Jensen (University of Copenhagen) for providing components for conjugation, and Eva Steel (University of Edinburgh) for assistance with assembling the CyanoGate kit. RV, AJM, BW and CJH were funded by the PHYCONET Biotechnology and Biological Sciences Research Council (BBSRC) Network in Industrial Biotechnology and Bioenergy (NIBB) and the Industrial Biotechnology Innovation Centre (IBioIC). GARG was funded by a BBSRC EASTBIO CASE PhD studentship (grant number BB/M010996/1). BW acknowledges funding support by UK BBSRC grant (BB/N007212/1) and Leverhulme Trust research grant (RPG-2015-445). AAS was funded by a Consejo Nacional de Ciencia y Tecnología (CONACYT) PhD studentship. KV was supported by a CSIRO Synthetic Biology Future Science Fellowship.

FIGURE LEGENDS

Figure 1. Adaptation of the Plant Golden Gate MoClo level 0 syntax for generating level 1 assemblies for transfer to Level T. (A) The format for a level 0 MoClo acceptor vector with the part bordered by two BsaI sites. (B) Typical level 0 parts from the Plant MoClo kit (38), where parts of the same type are bordered by the same pair of fusion sites (for each fusion site, only the sequence of the top strand is shown). Note that the parts are not drawn to scale. (C, D) The syntax of the Plant MoClo kit was adapted to generate level 0 parts for engineering marked and unmarked cyanobacterial mutant strains (20). (E-I) To generate knock in mutants, short linker parts (30 bp) were constructed to allow assembly of individual flanking sequences, or marker cassettes (Ab^R or *sacB*) in level 1 vectors for subsequent assembly in level T. (J, K) Parts required for generating synthetic srRNA or CRISPRi level 1 constructs. See **Supplementary Information S3 and S4** for workflows. Abbreviations: 3U+Ter, 3'UTR and terminator; Ab^R, antibiotic resistance cassette; Ab^R DOWN LINKER, short sequence (~30 bp) to provide CGCT overhang; Ab^R UP LINKER, short sequence (~30 bp) to provide GAGG overhang; CDS2(stop), coding sequence with a stop codon; DOWN FLANK, flanking sequence downstream of target site; DOWN FLANK LINKER, short sequence (~30 bp) to provide GGAG overhang; Prom+5U, promoter and 5' UTR; Prom TSS, promoter transcription start site; *sacB*, levansucrase expression cassette; *sacB* UP LINKER, short sequence (~30 bp) to provide GAGG overhang; sgRNA, single guide RNA; SP, signal peptide; srRNA, small regulatory RNA; UP FLANK, flanking sequence upstream of target site; UP FLANK LINKER, short sequence (~30 bp) to provide CGCT overhang; UNMARK LINKER, short sequence to bridge UP FLANK and DOWN FLANK.

Figure 2. Extension of the Plant Golden Gate MoClo Assembly Standard for cyanobacterial transformation. Assembly relies one of two Type IIS restriction endonuclease enzymes (*Bsa*I or *Bpi*I). Domesticated level 0 parts are assembled into level 1 vectors. Up to seven level 1 modules can be assembled directly into a level T cyanobacterial transformation vector, which consist of two sub-types (either a replicative or an integrative vector). Alternatively, larger vectors with more modules can be built by assembling level 1 modules into level M, and then cycling assembly between level M and level P, and finally transferred from Level P to level T. Antibiotic selection markers are shown for each level. Level T vectors are supplied with internal antibiotic selection markers (shown), but additional selection markers could be included from level 1 modules as required. See **Supplementary Table S1** and **Supplementary Information S1** for full list and maps of level T acceptor vectors.

Figure 3. Generating knock out mutants in cyanobacteria. (A) Assembled level T vector *cpcBA*-M (see Fig. 1C) targeting the *cpcBA* promoter and operon (3,563 bp) to generate a marked Δ*cpcBA* "Olive" mutant in *Synechocystis* sp. PCC 6803. Following transformation and segregation on kanamycin (*ca*. 3 months), a segregated marked mutant was isolated (WT band is 3,925 bp, marked mutant band is 5,503 bp, 1kb DNA ladder (NEB) is shown). (B) Assembled level T vector *cpcBA*-UM (see Fig. 1D) for generating an unmarked Δ*cpcBA* mutant. Following transformation and segregation on sucrose (*ca*. 2 weeks), an unmarked mutant was isolated (unmarked band is 425 bp). (C) Liquid cultures of WT, marked and unmarked Olive mutants. (D) Spectrum showing the absorbance of the unmarked Olive mutant and WT cultures after 72 hr of growth. Values are the average of four biological replicates ± SE and are standardised to 750 nm.

Figure 4. Generating knock in mutants in cyanobacteria. (A) Assembly of level 1 modules *cpcBA*-UF (see Fig. 1E) in the level 1, position 1 acceptor (L1P1), P_{cpc560}-eYFP-T_{rrnB} (see Fig. 1G) in L1P2 and *cpcBA*-DF (see Fig. 1F) in L1P3. (B) Transfer of level 1 assemblies to level T vector *cpcBA*-eYFP for generating an unmarked Δ*cpcBA* mutant carrying an eYFP expression cassette. Following transformation and segregation on sucrose (ca. 3 weeks), an unmarked eYFP mutant was isolated (1,771 bp). (C) Fluorescence values are the means ± SE of four biological replicates, where each replicate represents the median measurements of 10,000 cells.

Figure 5. Expression levels of cyanobacterial promoters in *Synechocystis* and UTEX 2973. (A) Structure of the cyanobacterial promoters adapted for the CyanoGate kit. Regions of P_{cpc560} shown are the upstream transcription factor binding sites (TFBSs) (-556 to -381 bp), middle region (-380 to -181 bp), and the downstream TFBSs, ribosome binding site (RBS) and spacer (-180 to -5 bp) (B) Expression levels of eYFP driven by promoters in *Synechocystis* and UTEX 2973 calculated from measurements taken from 10,000 individual cells. Values are the means ± SE from at least four biological replicates after 48 hr of growth (average OD_{750} values for *Synechocystis* and UTEX 2973 cultures were 3.5 ± 0.2 and 3.6 ± 0.2, respectively). See **Supplementary Figure S2** for more info.

Figure 6. Expression levels of heterologous and synthetic promoters in *Synechocystis* and UTEX 2973. (A) Structure and alignment of eight new synthetic promoters derived from the

BioBricks BBa_J23119 library and P_{trc10} promoter design (18). (B) Expression levels of eYFP driven by promoters in *Synechocystis* and UTEX 2973 calculated from measurements taken from 10,000 individual cells. (C) Correlation analysis of expression levels of synthetic promoters tested in *Synechocystis* and UTEX 2973. The coefficient of determination (R^2) is shown for the J23119 library (red), new synthetic promoters (pink) and *trc* variants (dark red). Values are the means ± SE from at least four biological replicates after 48 hr of growth (average OD_{750} values for *Synechocystis* and UTEX 2973 cultures were 3.5 ± 0.2 and 3.6 ± 0.2, respectively). See **Supplementary Figure S2** for more info.

Figure 7. Protein expression levels in *Synechocystis* and UTEX 2973 cells. (A) Confocal images of WT strains and mutants expressing eYFP (fluorescence shown in yellow) driven by the J23119 promoter (bar = 10 µm). (B) Representative immunoblot of protein extracts (3 µg protein) from mutants with different promoter expression cassettes (as in **Fig. 6**) probed with an antibody against eYFP. The protein ladder band corresponds to 30 kDa. (E) Relative eYFP protein abundance relative to UTEX 2973 mutants carrying the J23119 expression cassette. (C-E) Cell density, protein content per cell and protein density per estimated cell volume for *Synechocystis* and UTEX 2973. Asterisks (*) indicate significant difference ($P < 0.05$) as determined by Student's *t*-tests. Values are the means ± SE of four biological replicates.

Figure 8. Growth and expression levels of eYFP with the RK2 replicative origin in *Synechocystis*. (A) Growth of strains carrying RK2 (vector pSEVA421-T-eYFP), RSF1010 (pPMQAK1-T-eYFP) or an empty pPMQAK1-T grown in appropriate antibiotics. Growth was measured as OD_{750} under a constant illumination of 100 µmol photons $m^{-2}s^{-1}$ at 30 °C. (B) Expression levels of eYFP after 48 hr of growth calculated from measurements taken from 10,000 individual cells. (C) Plasmid copy numbers per cell after 48 hr of growth. Letters indicating significant difference ($P < 0.05$) are shown, as determined by ANOVA followed by Tukey's HSD tests. Values are the means ± SE of four biological replicates.

Figure 9. Gene regulation system using CRISPRi in *Synechocystis*. (A) Four target regions were chosen as sgRNA protospacers to repress eYFP expression in Olive-eYFP (**Fig. 4**): 'CCAGGATGGGCACCACCC' (+31), 'ACTTCAGGGTCAGCTTGCCGT' (+118), 'AGGTGGTCACGAGGGTGGGCCA' (+171) and 'AGAAGTCGTGCTGCTTCATG' (+233). (B) eYFP fluorescence of Olive-eYFP expressing constructs carrying sgRNAs with and without dCas9 (representative of 10,000 individual cells). Untransformed Olive-eYFP and

the Olive mutant were used as controls. Letters indicating significant difference ($P < 0.05$) are shown, as determined by ANOVA followed by Tukey's HSD tests. Values are the means ± SE of four biological replicates.

REFERENCES

Abe, K., Sakai, Y., Nakashima, S., Araki, M., Yoshida, W., Sode, K., Ikebukuro, K. (2014) Design of riboregulators for control of cyanobacterial (*Synechocystis*) protein expression. *Biotechnol. Lett.*, **36**, 287–294.

Albers, S.C., Gallegos, V.A., Peebles, C.A.M. (2015) Engineering of genetic control tools in *Synechocystis* sp. PCC 6803 using rational design techniques. *J. Biotechnol.*, **216**, 36–46.

Andersson, C.R., Tsinoremas, N.F., Shelton, J., Lebedeva, N.V., Yarrow, J., Min, H., Golden, S.S. (2000) Application of bioluminescence to the study of circadian rhythms in cyanobacteria. *Methods Enzymol.*, **305**,527–542.

Andreou, A.I., Nakayama, N. (2018) Mobius Assembly: A versatile Golden-Gate framework towards universal DNA assembly. *PLoS ONE*, **13**, e0189892.

Armshaw, P., Carey, D., Sheahan, C., Pembroke, J.T. (2015) Utilising the native plasmid, pCA2.4, from the cyanobacterium *Synechocystis* sp. strain PCC6803 as a cloning site for enhanced product production. *Biotechnol. Biofuels* **8**, 201.

Behler, J., Vijay, D., Hess, W.R., Akhtar, M.K. (2018) CRISPR-based technologies for metabolic engineering in cyanobacteria. *Trends Biotechnol.*, pii: S0167-7799(18)30146-X.

Bhal, C.P., Wu, R.,Stawinsky, J. Narang, S.A. (1977) Minimal length of the lactose operator sequence for the specific recognition by the lactose repressor. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 966–970.

Blasi, B., Peca, L., Vass, I., Kós, P.B. (2012) Characterization of stress responses of heavy metal and metalloid inducible promoters in *Synechocystis* PCC6803. *J. Microbiol. Biotechnol.*, **22**, 166–169.

Blasina, A., Kittell, B.L., Toukdarian, A.E., Helinski, D.R. (1996) Copy-up mutants of the plasmid RK2 replication initiation protein are defective in coupling RK2 replication origins. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 3559–3564.

Bradley, R.W., Buck, M., Wang, B. (2016) Tools and principles for microbial gene circuit engineering. *J. Mol. Biol.*, **428**, 862–882.

Bradley, R.W., Wang, B. (2015) Designer cell signal processing circuits for biotechnology. *N. Biotechnol.*, **32**, 635–643.

Bustos S.A., Golden, S.S. (1992) Light-regulated expression of the PsbD gene family in *Synechococcus* sp. strain PCC 7942: evidence for the role of duplicated PsbD genes in cyanobacteria. *Mol. Gen. Genet.*, **232**, 221–230.

Castenholz, R.W. (2001) Phylum BX. Cyanobacteria. Oxygenic photosynthetic bacteria. In: Garrity,G., Boone,D.R., Castenholz,R.W. (eds) Bergey's manual of systematic bacteriology. Springer, New York, pp 473–599.

Camsund, D., Heidorn, T., and Lindblad, P. (2014) Design and analysis of LacI-repressed promoters and DNA-looping in a cyanobacterium. *J. Biol. Eng.*, **8**, e4.

Chambers, S., Kitney, R., Freemont, P. (2016) The Foundry: the DNA synthesis and construction Foundry at Imperial College. *Biochem. Soc. Trans.*, **44**, 687–688.

Chen, Y., Taton, A., Go, M., London, R.E., Pieper, L.M., Golden, S.S., Golden, J.W. (2016). Self-replicating shuttle vectors based on pANS, a small endogenous plasmid of the unicellular cyanobacterium *Synechococcus elongatus* PCC 7942. *Microbiology*, **162**, 2029–2041.

Clerico, E.M., Ditty,J.L., Golden, S.S. (2007) Specialized techniques for site-directed mutagenesis in cyanobacteria. *Methods Mol. Biol.*, **362**, 155–171.

Crozet, P., Navarro, F.J., Willmund, F., Mehrshahi, P., Bakowski, K., Lauersen, K.J., Pérez-Pérez, M.E., Auroy, P., Gorchs Rovira, A., Sauret-Gueto,S., Niemeyer, J., Spaniol, B., Theis, J., Trösch, R., Westrich, L.D., Vavitsas, K., Baier, T., Hübner, W., de Carpentier, F., Cassarini, M., Danon, A., Henri, J., Marchand, C.H., de Mia, M., Sarkissian, K., Baulcombe, D.C., Peltier, G., Crespo, J.L., Kruse, O., Jensen, P.E., Schroda, M., Smith, A.G., Lemaire, S.D. (2018) Birth of a photosynthetic chassis: A MoClo toolkit enabling synthetic biology in the microalga *Chlamydomonas reinhardtii*. *ACS Synth. Biol.*, **7**, 2074 –2086.

Dexter, J., Dziga, D., Lv, J., Zhu, J., Strzalka, W., Maksylewicz, A., Maroszek, M, Marek, S., Fu, P. (2018). Heterologous expression of mlrA in a photoautotrophic host – Engineering cyanobacteria to degrade microcystins. *Environ. Pollut.*, **237**, 926–935.

Dobrin, A., Saxena, P., Fussenegger, M. (2016) Synthetic biology: applying biological circuits beyond novel therapies. *Integr. Biol.*, **8**, 409–430.

Doud, D.F.R., Holmes, E.C.,Richter, H., Molitor, B., Jander, G., Angenent, L.T. (2017) Metabolic engineering of Rhodopseudomonas palustris for the obligate reduction of n-butyrate to n-butanol. *Biotechnol. Biofuels.*, **10**, 178.

Ducat, D.C., Way, J.C., Silver, P.A. (2010) Engineering cyanobacteria to generate high-value products. *Trends Biotechnol.*, **29**, 95–103.

Economou, C., Wannathong, T., Szaub, J., Purton, S. (2014) A simple, low-cost method for chloroplast transformation of the green alga *Chlamydomonas reinhardtii*. *Methods Mol. Biol.*, **1132**, 401–411.

Engler, C., Youles, M., Gruetzner, R., Ehnert, T-M., Werner, S., Jones, J.D.G., Patron, N.J., Marillonnet, S. (2014) A Golden Gate modular cloning toolbox for plants. *ACS Synth. Biol.*, **3**, 839–843.

Englund, E., Liang, F., Lindberg, P. (2016) Evaluation of promoters and ribosome binding sites for biotechnological applications in the unicellular cyanobacterium *Synechocystis* sp. PCC 6803. *Sci. Rep.*, **6**, A36640.

Ferreira, E.A., Pacheco, C.C., Pinot, F., Pereira, J., Lamosa, P., Oliveira, P.,Kirov, B., Jaramillo, A., Tamagnini, P. (2018) Expanding the toolbox for Synechocystis sp. PCC 6803: validation of replicative vectors and characterization of a novel set of promoters. *Synth. Biol.*, **3**, ysy014.

Flombaum, P., Gallegos, J.L., Gordillo, R.A., Rincón, J., Zabala, L.L., Jiao, N., Karl, D.M., Li, W.K.W., Lomas, M.W., Veneziano, D., Vera, C.S., Vrugt, J.A., Martiny, A.C. (2013) Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 9824–9829.

Frey, J., Bagdasarian, M.M., Bagdasarian, M. (1992) Replication and copy number control of the broad-host-range plasmid RSF1010. *Gene*, **113**, 101–106.

Gao, X., Gao, F., Liu, D., Zhang, H., Nie, X., Yang, C. (2016) Engineering the methylerythritol phosphate pathway in cyanobacteria for photosynthetic isoprene production from CO_2. *Energy Environ. Sci.*, **9**, 1400–1411.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison C.A., Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, **6**, 343–345.

Gordon, G.C., Korosh, T.C., Cameron, J.C., Markley, A.L., Begemann, M.B., Pfleger, B.F. (2016) CRISPR interference as a titratable, trans-acting regulatory tool for metabolic engineering in the cyanobacterium *Synechococcus* sp. strain PCC 7002. *Metab. Eng.*, **38**, 170–179.

Guerrero, F., Carbonell, V., Cossu, M., Correddu, D., Jones, P.R. (2012) Ethylene synthesis and regulated expression of recombinant protein in *Synechocystis* sp. PCC 6803. *PLoS ONE*, **7**, e50470.

Immethun, C.M., DeLorenzo, D.M., Focht, C.M., Gupta, D., Johnson, C.B., Moon, T.S. (2017) Physical, chemical, and metabolic state sensors expand the synthetic biology toolbox for *Synechocystis* sp. PCC 6803. *Biotechnol Bioeng.*, **114**, 1561–1569.

Heidorn, T., Camsund, D., Huang, H., Lindberg, P., Oliveira, P., Stensjö, K., Lindblad, P. (2011) Synthetic biology in cyanobacteria engineering and analyzing novel functions. *Methods Enzymol.*, **497**, 539–579.

Hendry, J.I., Prasannan, C.B., Joshi, A., Dasgupta, S., Wangikar, P.P. (2016) Metabolic model of *Synechococcus* sp. PCC 7002: Prediction of flux distribution and network modification for enhanced biofuel production. *Bioresour Technol.*, **213**, 190–197.

Heyduk, E., Heyduk, T. (2018) DNA template sequence control of bacterial RNA polymerase escape from the promoter. *Nucleic Acids Res.*, **46**, 4469–4486.

Higo, A., Isu, A., Fukaya, Y., Ehira, S., Hisabori, T. (2018) Application of CRISPR interference for metabolic engineering of the heterocyst-forming multicellular cyanobacterium Anabaena sp. PCC 7120. *Plant Cell Physiol.*, **59**, 119–127.

Higo, A., Isu, A., Fukaya, Y., Hisabori, T. (2016) Efficient gene induction and endogenous gene repression systems for the filamentous cyanobacterium *Anabaena* sp. PCC 7120.

Huang, C.H., Shen, C.R., Li, H., Sung, L.Y., Wu, M.Y., Hu, Y.C. (2016) CRISPR interference (CRISPRi) for gene regulation and succinate production in cyanobacterium S. elongatus PCC 7942. *Microb. Cell Fact.*, **15**, 196.

Huang, H., Camsund, D., Lindblad, P., Heidorn, T. (2010) Design and characterization of molecular tools for a Synthetic Biology approach towards developing cyanobacterial biotechnology. *Nucleic Acids Res.*, **38**, 2577–2593.

Huang, H., Lindblad, P. (2013) Wide-dynamic-range promoters engineered for cyanobacteria. *J. Biol. Eng.*, **7**, 10.

Jusiak, B., Cleto, S., Perez-Piñera, P., Lu, T.K. (2016) Engineering synthetic gene circuits in living cells with CRISPR technology. *Trends Biotechnol.*, **34**, 535–547.

Keeling,P.J. (2003) Diversity and evolutionary history of plastids and their hosts. *Am. J. Bot.*, **10**, 1481–1493.

Kelly, C.L., Hitchcock, A., Torres-Méndez, A., Heap, J.T. (2018) A rhamnose-inducible system for precise and temporal control of gene expression in cyanobacteria. *ACS Synth. Biol.*, **7**, 1056–1066.

Kikuchi, H., Aichi, M., Suzuki, I., Omato, T. (1996) Positive regulation by nitrite of the nitrate assimilation operon in the cyanobacteria *Synechococcus* sp. strain PCC 7942 and *Plectonema boryanum*. *J. Bacteriol.*, **178**, 5822–5825.

Kim, W.J., Lee, S.M., Um, Y., Sim, S.J., Woo, H.M. (2017) Development of SyneBrick vectors as a synthetic biology platform for gene expression in *Synechococcus elongatus* PCC 7942. *Front. Plant Sci.*, **8**, A293.

Kirst, H., Formighieri, C., Melis, A. (2014) Maximizing photosynthetic efficiency and culture productivity in cyanobacteria upon minimizing the phycobilisome light-harvesting antenna size. *Biochim. Biophys. Acta*, **1837**, 1653–1664.

Knoop, H., Gründel, M., Zilliges, Y., Lehmann, R., Hoffmann, S., Lockau, W., Steuer, R. (2013) Flux Balance Analysis of Cyanobacterial Metabolism: The Metabolic Network of *Synechocystis* sp. PCC 6803. *PLoS Comput. Biol.*, **9**, e1003081.

Kulkarni, D.R., Golden, S.S. (1997). mRNA stability is regulated by a coding-region element and the unique 5′ untranslated leader sequences of the three Synechococcus psbA transcripts. *Mol. Microbiol.*, **24**, 1131–1142.

Kunert, A., Vinnemeier, J., Erdmann, N., Hagemann, M. (2003) Repression by Fur is not the main mechanism controlling the iron-inducible isiAB operon in the cyanobacterium *Synechocystis* sp. PCC 6803. *FEMS Microbiol. Lett.*, **227**, 255–262.

Lea-Smith, D.J., Bombelli, P., Dennis, J.S., Scott, S.A., Smith, A.G. Howe, C.J. (2014) Phycobilisome-deficient strains of *Synechocystis* sp. PCC 6803 have reduced size and require carbon-limiting conditions to exhibit enhanced productivity. *Plant Physiol.*, **165**, 705−714.

Lea-Smith, D.J., Ross, N., Zori, M., Bendall, D.S., Dennis, J.S., Scott, S.A., Smith, A.G., Howe, C.J. (2013) Thylakoid terminal oxidases are essential for the cyanobacterium *Synechocystis* sp. PCC 6803 to survive rapidly changing light intensities. *Plant Physiol.*, **162**, 484–495.

Lea-Smith, D.J., Vasudevan, R., Howe, C.J. (2016) Generation of marked and markerless mutants in model cyanobacterial species. *J. Vis. Exp.*, **111**, e54001.

Leonard, S.P., Perutka, J., Powell, J.E., Geng, P., Richhart, D.D., Byrom, M., Kar, S., Davies, B.W., Ellington, A.D., Moran, N.A., Barrick, J.E. (2018) Genetic engineering of bee gut microbiome bacteria with a toolkit for modular assembly of broad-host-range plasmids. *ACS Synth. Biol.*, **7**, 1279–1290.

Li, L., Jaing, W., Lu, Y. (2018) A modified Gibson Assembly method for cloning large DNA fragments with high GC contents. Methods Mol. Biol., **1671**, 203-209.

Li, S., Sun, T., Xu, C., Chen, L., Zhang, W. (2018) Development and optimization of genetic toolboxes for a fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Metab. Eng.*, **48**, 163–174.

Liberton, M., Chrisler, W.B., Nicora, C.D., Moore, R.J., Smith, R.,D., Koppenaal, D.W., Pakrasi, H.B., Jacobs, J.M. (2017) Phycobilisome truncation causes widespread proteome changes in *Synechocystis* sp. PCC 6803. *PLoS ONE*, **12**, e0173251.

Lindberg, P., Park, S., Melis, A. (2010) Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metab. Eng.*, **12**, 70–79.

Luan, G., Lu, X. (2018) Tailoring cyanobacterial cell factory for improved industrial properties. *Biotechnol. Adv.*, **36**, 430–442.

Liu, Q., Schumacher, J., Wan, X., Lou, C., Wang, B. (2018) Orthogonality and burdens of heterologous AND Gate gene circuits in E. coli. *ACS Synth. Biol.*, **7**, 553–564.

Liu,D., Pakrasi, B.H. (2018) Exploring native genetic elements as plug-in tools for synthetic biology in the cyanobacterium *Synechocystis* sp. PCC 6803. *Microb. Cell Fact.*, **17**,48.

Lou, W., Tan, X., Song, K., Zhang, S., Luan, G., Li, C., Lu, X. (2018) A specific single nucleotide polymorphism in the ATP synthase gene significantly improves environmental stress tolerance of *Synechococcus elongatus* PCC 7942. *Appl. Environ. Microbiol.*, **84**, e01222-18.

Madsen, M.A., Semerdzhiev, S., Amtmann, A., Tonon, T. (2018) Engineering mannitol biosynthesis in *Escherichia coli* and *Synechococcus* sp. PCC 7002 using a green algal fusion protein. *ACS Synth. Biol.*, **7**, 2833–2840.

Markley, A.L., Begemann, M.B., Clarke, R.E., Gordon, G.C., Pflege, B.F. (2015) Synthetic biology toolbox for controlling gene expression in the cyanobacterium *Synechococcus* sp. strain PCC 7002. *ACS Synth Biol.*, **4**, 595−603.

Marraccini, P., Bulteau, S., Cassier-Chauvat, C., Mermet-Bouvier, P., Chauvat, F. (1993) A conjugative plasmid vector for promoter analysis in several cyanobacteria of the genera *Synechococcus* and *Synechocystis*. *Plant Mol. Biol.*, **23**, 905–909.

McCormick, A.J., Bombelli, P., Bradley, R.W., Thorne, R., Wenzele, T., Howe, C.J. (2015) Biophotovoltaics: oxygenic photosynthetic organisms in the world of bioelectrochemical systems. *Energy Environ. Sci.*, **8**, 1092–1109.

Medford. J.I., Prasad. A. (2016) Towards programmable plant genetic circuits. *Plant J.*, **87**, 139–148.

Mermet-Bouvier, P., Cassier-Chauvat, C., Marraccini, P., Chauvat, F. (1993) Transfer and replication of RSF1010-derived plasmids in several cyanobacteria of the general *Synechocystis* and *Synechococcus*. *Curr. Microbiol.* **27**, 323–327.

Mohammadi, R., Fallah-Mehrabadi, J., Bidkhori, G., Zahiri, J., Javad Niroomand, M., Masoudi-Nejad, A. (2016) A systems biology approach to reconcile metabolic network models with application to *Synechocystis* sp. PCC 6803 for biofuel production. *Mol. Biosyst.*, **12**, 2552 –2561.

Moore, S.J., Lai, H-E., Kelwick, R.J.R., Chee, S.M., Bell, D.J., Polizzi, K.M., Freemont, P.S. (2016) EcoFlex: a multifunctional MoClo kit for *E. coli* synthetic biology. *ACS Synth. Biol.*, **5**, 1059−1069.

Mueller, T.J., Ungerer, J.L., Pakrasi, H.B., Maranas, C.D. (2017) Identifying the metabolic differences of a fast-growth phenotype in Synechococcus UTEX 2973. *Sci. Rep.*, **7**,41569.

Munch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., Jahn, D. (2005) Genome analysis Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinforma. Appl. Note* **21**, 4187–4189.

Mutalik, V.K., Guimaraes, J.C., Cambray, G., Mai, Q.A., Christoffersen, M.J., Martin, L., Yu, A., Lam, C., Rodriguez, C., Bennett, G., Keasling, J.D., Endy, D., Arkin, A.P. (2013) Quantitative estimation of activity and quality for collections of functional genetic elements. *Nat. Methods* **10**, 347–353.

Na, D., Yoo, S.M., Chung, H., Park, H., Park, J.H., Lee, S.Y. (2013) Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. *Nat. Biotechnol.*, **31**, 170–174.

Niederholtmeyer, H., Wolfstädter, B.T., Savage, D.F., Silver, P.A., Way, J.C. (2010) Engineering cyanobacteria to synthesize and export hydrophilic products. *Appl. Environ. Microbiol.*, **76**, 3462–3466.

Nielsen, A.Z., Mellor, S.B., Vavitsas, K., Wlodarczyk, A.J., Gnanasekaran, T., de Jesus, M.P.R.H., King, B.C. Bakowski, K., Jensen, P.E. (2016) Extending the biosynthetic repertoires of cyanobacteria and chloroplasts. *Plant J.*, **87**, 87–102.

Nielsen, J., Keasling, J.D. (2016) Engineering cellular metabolism. *Cell*, **164**, 1185–1197.

Ng, A.H., Berla, B.M., Pakrasi, H.B. (2015) Fine-tuning of photoautotrophic protein production by combining promoters and neutral sites in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Appl. Environ. Microbiol.* **81**, 6857–6863.

Ng, W.O., Zentella, R., Wang, Y., Taylor, J.S., Pakrasi, H.B. (2000) PhrA, the major photoreactivating factor in the cyanobacterium *Synechocystis* sp. strain PCC 6803 codes for a cyclobutane-pyrimidine-dimer-specific DNA photolyase. *Arch. Microbiol.*, **173**, 412–417.

Ordon, J., Gantner, J., Kemna, J., Schwalgun, L., Reschke, M., Streubel, J., Boch, J., Stuttmann, J. (2017) Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit. *Plant J.*, **89**:155–168.

Patron, N.J., Orzaez, D., Marillonnet, S., Warzecha, H., Matthewman, C., Youles, M., Raitskin, O., Leveau, A., Farré, G., Rogers, C., Smith, A., Hibberd, J., Webb, A.A. *et al.* (2015) Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *New Phytol.*, **208**,13–19.

Peca, L., Kós, P.B., Máté, Z., Farsang, A., Vass, I. (2008) Construction of bioluminescent cyanobacterial reporter strains for detection of nickel, cobalt and zinc. *FEMS Microbiol. Lett.*, **289**, 258–264.

Pinto, F., Pacheco, C.,C., Ferreira, D., Moradas-Ferreira, P., Tamagnini, P. (2012) Selection of suitable reference genes for RT-qPCR analyses in cyanobacteria. *PLoS ONE*, **7**, e34983.

Pinto, F., Pacheco, C.,C., Oliveira, P., Montagud, A., Landels, A., Couto, N., Wright, P.,C., Urchueguía, J.,F., Tamagnini, P. (2015) Improving a *Synechocystis*-based photoautotrophic

chassis through systematic genome mapping and validation of neutral sites. *DNA Res.*,**22**, 425–437.

Pye, C.R., Bertin, M.J., Lokey, R.S., Gerwick, W.H., Linington, R.G. (2017) Retrospective analysis of natural products provides insights for future discovery trends. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 5601–5606.

Qi, Q., Hao, M., Ng, W., Slater, S.C., Baszis, S.R., Weiss, J.D., Valentin, H.E. (2005) Application of the *Synechococcus nirA* promoter to establish an inducible expression system for engineering the *Synechocystis* tocopherol pathway. *Appl. Environ. Microbiol.*, **71**, 5678–5684.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S. Arkin, A.P., Lim, W.A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, **152**, 1173–1183.

Rae, B.D., Long, B., Förster, B., Nguyen, N.D., Velanis, C.N., Atkinson, N., Hee, W., Mukherjee, B., Price, G.D., McCormick, A.J. (2017) Progress and challenges of engineering a biophysical carbon dioxide concentrating mechanism into higher plants. *J. Ex. Bot.*, **68**, 3717–3737.

Ramey, C,.J., Barón-Sola, Á., Aucoin, H.R., Boyle, N.R. (2015) Genome Engineering in Cyanobacteria: Where We Are and Where We Need To Go. *ACS Synth. Biol.*, **4**, 1186–1196.

Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., Stanier, R.Y. (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J. Gen. Microbiol.*, **111**, 1–61.

Ruffing, A. M., Jensen, T. J., Strickland, L. M. (2016) Genetic tools for advancement of *Synechococcus* sp. PCC 7002 as a cyanobacterial chassis. *Microb. Cell Fact.*, **15**, 1–14.

Saar, K.L., Bombelli, P., Lea-Smith, D.J., Call, T., Aro, E-M., Müller, T., Howe, C.J., Knowles, T.P.J. (2018) Enhancing power density of biophotovoltaics by decoupling storage and power delivery. *Nat. Energy*, **3**, 75–81.

Sakai, Y., Abe, K., Nakashima, S., Ellinger, J.J., Ferri, S., Sode, K., Ikebukuro, K. (2015) Scaffold-fused riboregulators for enhanced gene activation in *Synechocystis* sp. PCC 6803. *MicrobiologyOpen*, **4**, 533–540.

Sarrion-Perdigones, A., Vazquez-Vilar, M., Palací, J., Castelijns, B., Forment, J., Ziarsolo, P., Blanca, J., Granell, A., Orzaez, D. (2013) GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiol.*, **162**, 1618–1631.

Shirai,T., Osanai,T. Kondo,A. (2016) Designing intracellular metabolism for production of target compounds by introducing a heterologous metabolic reaction based on a *Synechocystis* sp. 6803 genome-scale model. *Microb. Cell Fact.*, **15**, 13.

Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., de Las Heras, A., Páez-Espino, A.D., Durante-Rodríguez, G., Kim, J., Nikel, P.I., Platero, R., de Lorenzo, V. (2013) The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* **41**, D666–675.

Smanski, M.J., Zhou, H., Claesen, J., Shen, B., Fischbach, M.A., Voigt, C.A. (2016) Synthetic biology to access and expand nature's chemical diversity. *Nat. Rev. Microbiol.*, **14**, 135–149.

Stensjö, K., Vavitsas, K., Tyystjärvi, T. (2017) Harnessing transcription for bioproduction in cyanobacteria. *Physiol. Plant*,**162**, 148–155.

Stucken, K., Ilhan, J., Roettger, M., Dagan, T. Martin, W.F. (2012) Transformation and conjugal transfer of foreign genes into the filamentous multicellular cyanobacteria (subsection V) *Fischerella* and *Chlorogloeopsis*. *Curr. Microbiol.*, **65**, 552–560.

Sun, T., Li, S., Song, X., Diao, J., Chen, L., Zhang, W. (2018) Toolboxes for cyanobacteria: Recent advances and future direction. *Biotechnol Adv*., **36**, 1293–1307.

Tan, X., Hou, S., Song, K., Georg, J., Klähn, S., Lu, X., Hess, W.R. (2017) The primary transcriptome of the fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Biotechnol Biofuels.*, **11**, 218.

Tan, X., Yao, L., Gao, Q., Wang, W., Qi, F., Lu, X. (2011) Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria. *Metab. Eng.*, **13**, 169–176.

Taton, A.,U., Ma, A.T., Ota, M., Golden, S.S., Golden, J.W. (2017) NOT gate genetic circuits to control gene expression in cyanobacteria. *ACS Synth. Biol*., **6**, 2175−2182.

Taton, A.,U., Unglaub, F. Wright, N.E., Zeng, W.Y., Paz-Yepes, J., Brahamsha, B., Palenik, B. Peterson, T.C., Haerizadeh, F., Golden, S.S., Golden, J.W. (2014) Broad-host-range vector system for synthetic biology and biotechnology in cyanobacteria. *Nucleic Acids Res.*, **42**, e136.

Tsinoremas, N.F., Kutach, A.K., Strayer, C.A., Golden, S.S. (1994) Efficient gene transfer in *Synechococcus* sp. strains PCC 7942 and PCC 6301 by interspecies conjugation and chromosomal recombination. *J. Bacteriol.*, **176**, 6764–6768.

Ungerer, J., Lin, P-C., Chen, H-Y., Pakrasi, H.B. (2018a) Adjustments to photosystem stoichiometry and electron transfer proteins are key to the remarkably fast growth of the cyanobacterium *Synechococcus elongatus* UTEX 2973. *mBio*, **9**, e02327–17.

Ungerer, J., Pakrasi, H.B. (2016) Cpf1 Is a versatile tool for CRISPR genome editing across diverse species of cyanobacteria. *Sci. Rep.*, **6**, e39681.

Ungerer, J., Wendt K.E., Hendry J.I., Marans C.D., Pakrasi, H.B. (2018b) Comparative genomics reveals the molecular determinants of rapid growth of the cyanobacterium Synechococcus elongatus UTEX 2973. *Proc. Natl. Acad. Sci. U.S.A.*, doi/10.1073/pnas.1814912115.

van de Meene, A.M., Hohmann-Marriott, M.F., Vermaas, W.F., Roberson, R.W. (2006) The three-dimensional structure of the cyanobacterium *Synechocystis* sp. PCC 6803. *Arch Microbiol.*, **184**,259–270.

Vazquez-Vilar M., Orzaez, D., Patron, N. (2018) DNA assembly standards: Setting the low-level programming code for plant biotechnology. *Plant Sci.*, **273**, 33–41.

Vioque, A. (2007) Transformation of Cyanobacteria. In: León,R., Galván,A., Fernández,E. (eds) Transgenic Microalgae as Green Cell Factories. Advances in Experimental Medicine and Biology, vol 616. Springer, New York, pp 12–22.

Vogel, A.I.M., Lale,R., Hohmann-Marriott, M.F. (2017) Streamlining recombination-mediated genetic engineering by validating three neutral integration sites in *Synechococcus* sp. PCC 7002. *J. Biol. Eng.*, **11**, 19.

Wang, B., Eckert, C., Maness, P-C., Yu, J. (2018) A genetic toolbox for modulating the expression of heterologous genes in the cyanobacterium *Synechocystis* sp. PCC 6803. *ACS Synth. Biol.*, **19**, 276–286.

Watanabe S., Noda A., Ohbayashi R., Uchioke K., Kurihara A., Nakatake S., Morioka S., Kanesaki Y., Chibazakura T., Yoshikawa H. (2018) ParA-like protein influences the distribution of multi-copy chromosomes in cyanobacterium Synechococcus elongatus PCC 7942. *Microbiology*, **164**, 45–56.

Wendt, K.E., Ungerer, J., Cobb,R .E., Zhao, H., Pakrasi, H.B. (2016) CRISPR/Cas9 mediated targeted mutagenesis of the fast growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Microb. Cell. Fact.*, **15**, e115

Werner, S., Engler, C., Weber, E., Gruetzner, R., Marillonnet, S. (2012) Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. *Bioeng. Bugs.* **3**, 38–43.

Wlodarczyk, A., Gnanasekaran, T., Nielsen, A.Z., Zulu, N.N., Mellor, S.B., Luckner, M., Thøfner, J.F.B, Olsen, C.E., Mottawie, M.S., Burow, M., Pribil, M., Feussner, I., Møller, B.L., Jensen, P.E. (2016) Metabolic engineering of light-driven cytochrome P450 dependent pathways into *Synechocystis* sp. PCC 6803. *Metab. Eng.*, **33**, 1–11.

Xiao, A., Cheng, Z., Kong, L., Zhu, Z., Lin, S., Gao, G., Zhang, B. (2014) CasOT: a genome-wide Cas9/gRNA off-target searching tool. *Bioinformatics*, **30**, 1180–1182.

Xu, Y., Alvey, R.M., Byrne, P.O., Graham, J.E., Shen, G., Bryant, D. A. (2011) Expression of genes in cyanobacteria: adaptation of endogenous plasmids as platforms for high-level gene expression in *Synechococcus* sp. PCC 7002. *Methods Mol. Biol.*, **684**, 273–293.

Yao, L., Cengic, I., Anfelt, J., Hudson, E.P. (2015) Multiple gene repression in cyanobacteria using CRISPRi. *ACS Synth. Biol*., **5**, 207−212.

Yu, J., Liberton, M., Cliften, P.F., Head, R.D., Jacobs, J.M., Smith, R.D, Koppenaal, D.W., Brand, J.J., Pakrasi, H.B. (2015) *Synechococcus elongatus* UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO_2. *Sci. Rep.*, **5**, 8132.

Zess, E.K., Begemann, M.B., Pfleger, B.F. (2016) Construction of new synthetic biology tools for the control of gene expression in the cyanobacterium Synechococcus sp. strain PCC 7002. *Biotechnol. Bioeng.*, **113**, 424–432.

Zerulla, K., Ludt, K. Soppa J. (2016) The ploidy level of *Synechocystis* sp. PCC 6803 is highly variable and is influenced by growth phase and by chemical and physical external parameters. *Microbiology*, **162**, 730–739.

Zhou, J., Zhang, H., Meng, H., Zhu, H., Bao, G., Zhang, Y., Li, Y., Ma, Y. (2014) Discovery of a super-strong promoter enables efficient production of heterologous proteins in cyanobacteria. *Sci. Rep.*, **4**, 4500.

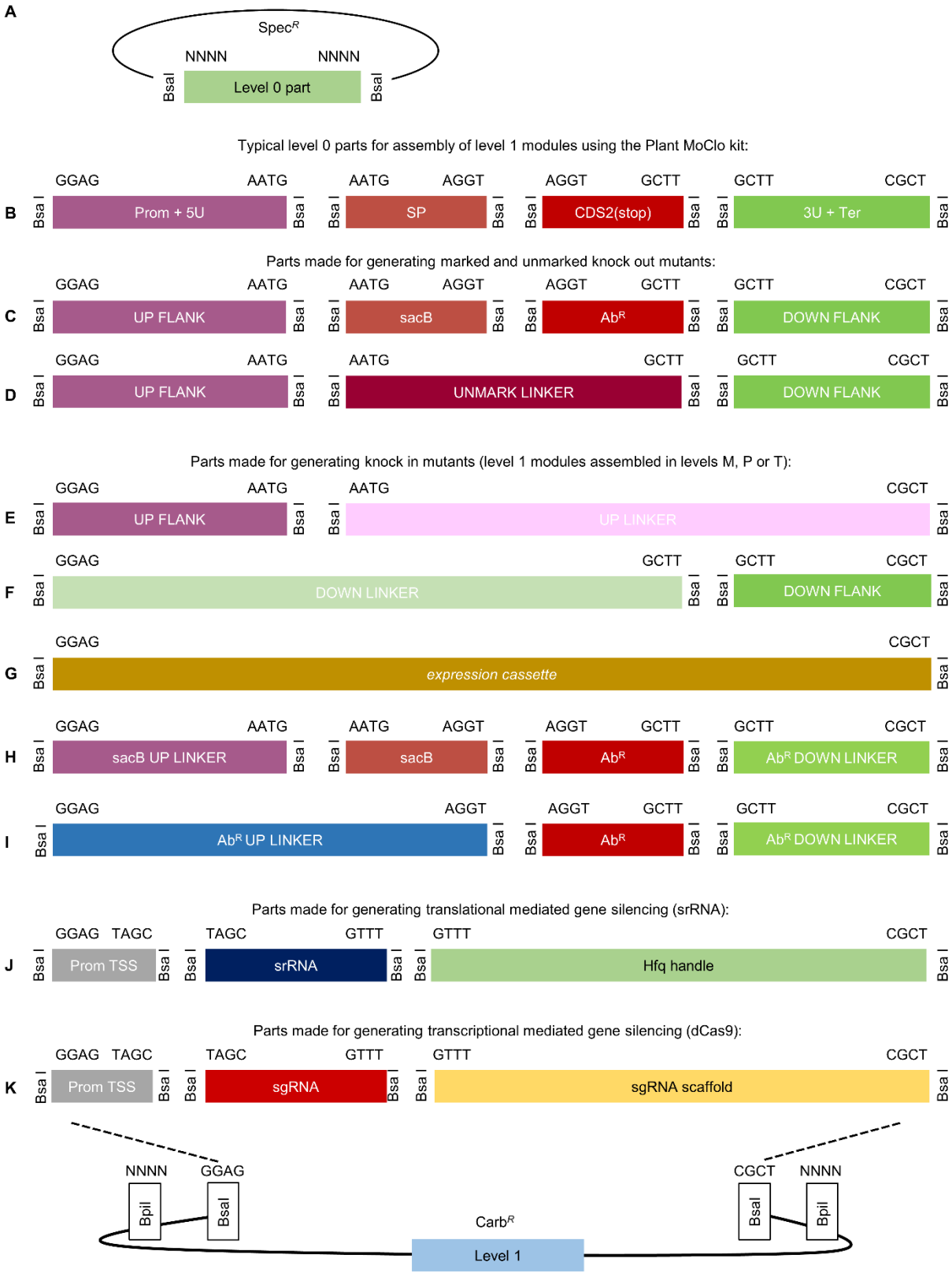


Figure 1. Adaptation of the Plant Golden Gate MoClo level 0 syntax for generating level 1 assemblies for transfer to Level T. (A) The format for a level 0 MoClo acceptor vector with the part bordered by two BsaI sites. (B) Typical level 0 parts from the Plant MoClo kit (38),

where parts of the same type are bordered by the same pair of fusion sites (for each fusion site, only the sequence of the top strand is shown). Note that the parts are not drawn to scale. (C, D) The syntax of the Plant MoClo kit was adapted to generate level 0 parts for engineering marked and unmarked cyanobacterial mutant strains (20). (E-I) To generate knock in mutants, short linker parts (30 bp) were constructed to allow assembly of individual flanking sequences, or marker cassettes (Ab^R or *sacB*) in level 1 vectors for subsequent assembly in level T. (J, K) Parts required for generating synthetic srRNA or CRISPRi level 1 constructs. See **Supplementary Information S3-S4** for workflows. Abbreviations: 3U+Ter, 3'UTR and terminator; Ab^R, antibiotic resistance cassette; Ab^R DOWN LINKER, short sequence (~30 bp) to provide CGCT overhang; Ab^R UP LINKER, short sequence (~30 bp) to provide GAGG overhang; CDS2(stop), coding sequence with a stop codon; DOWN FLANK, flanking sequence downstream of target site; DOWN FLANK LINKER, short sequence (~30 bp) to provide GGAG overhang; Prom+5U, promoter and 5' UTR; Prom TSS, promoter transcription start site; *sacB*, levansucrase expression cassette; *sacB* UP LINKER, short sequence (~30 bp) to provide GAGG overhang; sgRNA, single guide RNA; SP, signal peptide; srRNA, small regulatory RNA; UP FLANK, flanking sequence upstream of target site; UP FLANK LINKER, short sequence (~30 bp) to provide CGCT overhang; UNMARK LINKER, short sequence to bridge UP FLANK and DOWN FLANK.

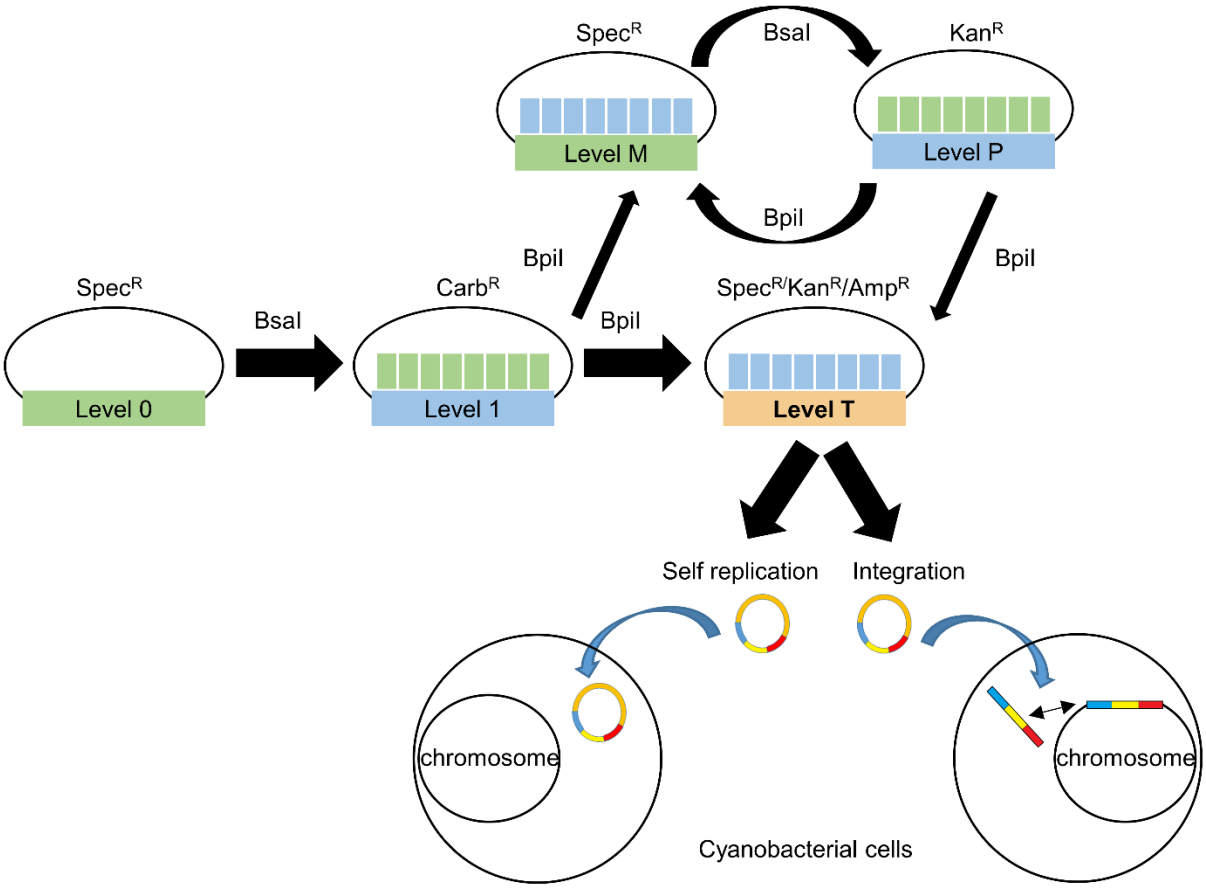


Figure 2. Extension of the Plant Golden Gate MoClo Assembly Standard for cyanobacterial transformation. Assembly relies one of two Type IIS restriction endonuclease enzymes (*Bsa*I or *Bpi*I). Domesticated level 0 parts are assembled into level 1 vectors. Up to seven level 1 modules can be assembled directly into a level T cyanobacterial transformation vector, which consist of two sub-types (either a replicative or an integrative vector). Alternatively, larger vectors with more modules can be built by assembling level 1 modules into level M, and then cycling assembly between level M and level P, and finally transferred from Level P to level T. Antibiotic selection markers are shown for each level. Level T vectors are supplied with internal antibiotic selection markers (shown), but additional selection markers could be included from level 1 modules as required. **Supplementary Table S1** and **Supplementary Information S1**.

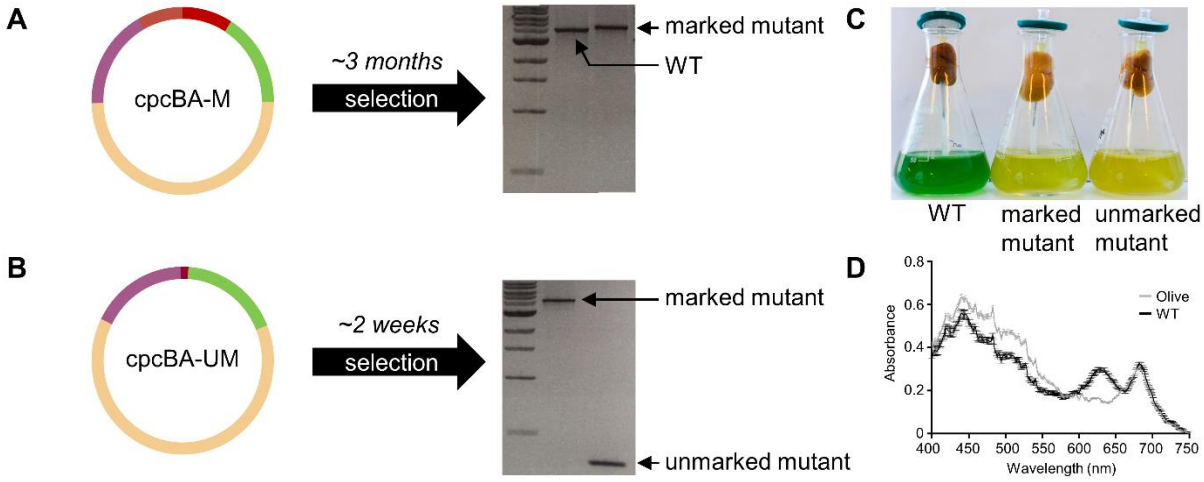


Figure 3. Generating knock out mutants in cyanobacteria. (A) Assembled level T vector *cpcBA*-M (see Fig. 1C) targeting the *cpcBA* promoter and operon (3,563 bp) to generate a marked Δ*cpcBA* "Olive" mutant in *Synechocystis* sp. PCC 6803. Following transformation and segregation on Kanamycin (*ca*. 3 months), a segregated marked mutant was isolated (WT band is 3,925 bp, marked mutant band is 5,503 bp, 1kb DNA ladder (NEB) is shown). (B) Assembled level T vector *cpcBA*-UM (see Fig. 1D) for generating an unmarked Δ*cpcBA* mutant. Following transformation and segregation on sucrose (*ca*. 2 weeks), an unmarked mutant was isolated (unmarked band is 425 bp). (C) Liquid cultures of WT, marked and unmarked Olive mutants. (D) Spectrum showing the absorbance of the unmarked Olive mutant and WT cultures after 72 hr of growth. Values are the average of four biological replicates ± SE and are standardised to 750 nm.

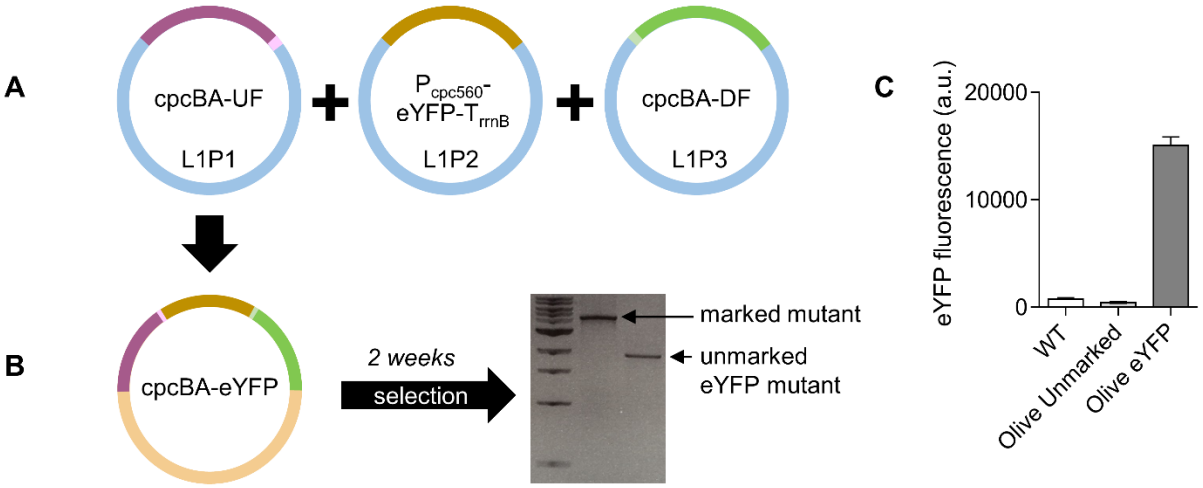


Figure 4. Generating knock in mutants in cyanobacteria. (A) Assembly of level 1 modules *cpcBA*-UF (see Fig. 1E) in the level 1, position 1 acceptor (L1P1), P_{cpc560}-eYFP-T_{rrnB} (see Fig. 1G) in L1P2 and *cpcBA*-DF (see Fig. 1F) in L1P3. (B) Transfer of level 1 assemblies to level T vector *cpcBA*-eYFP for generating an unmarked Δ*cpcBA* mutant carrying an eYFP expression cassette. Following transformation and segregation on sucrose (ca. 3 weeks), an unmarked eYFP mutant was isolated (1,771 bp). (C) Fluorescence values are the means ± SE of four biological replicates, where each replicate represents the median measurements of 10,000 cells.

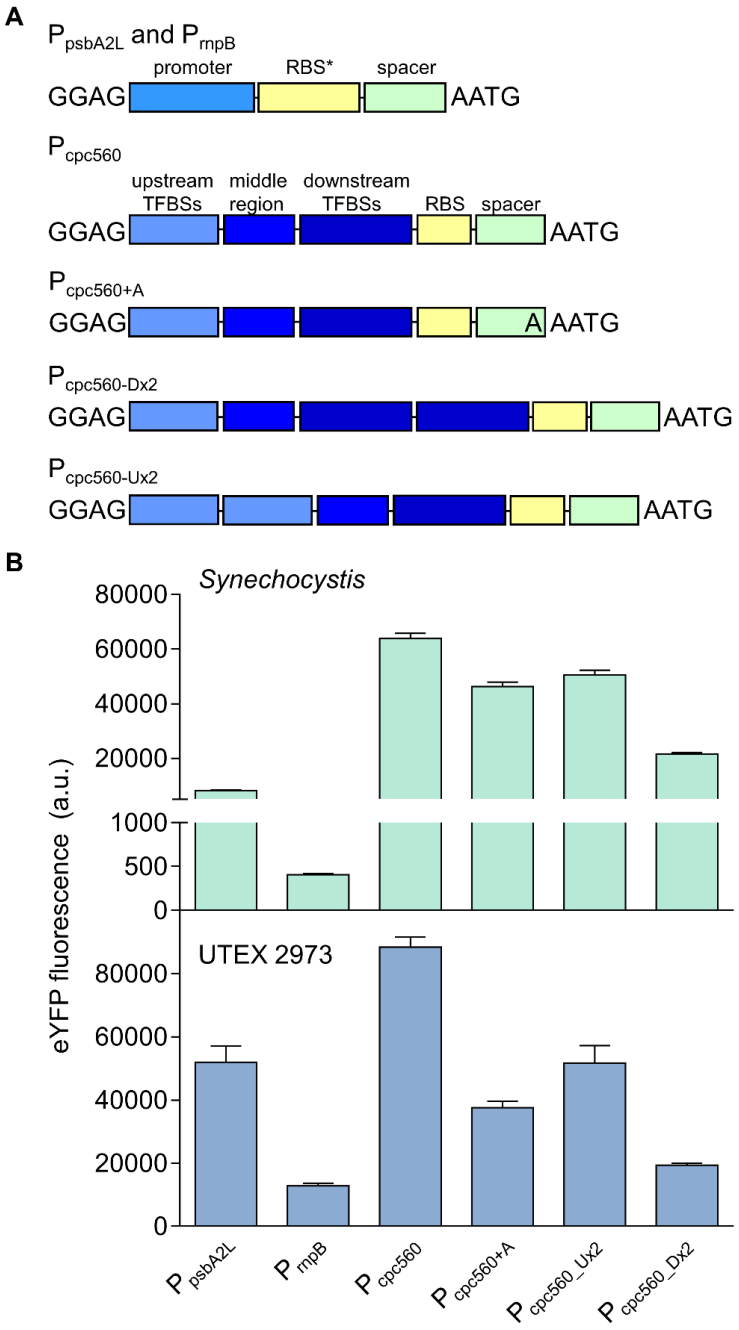


Figure 5. Expression levels of cyanobacterial promoters in *Synechocystis* and UTEX 2973. (A) Structure of the cyanobacterial promoters adapted for the CyanoGate kit. Regions of P_{cpc560} shown are the upstream transcription factor binding sites (TFBSs) (-556 to -381 bp), middle region (-380 to -181 bp), and the downstream TFBSs, ribosome binding site (RBS) and spacer (-180 to -5 bp) (B) Expression levels of eYFP driven by promoters in *Synechocystis* and UTEX 2973 calculated from measurements taken from 10,000 individual cells. Values are the means ± SE from at least four biological replicates after 48 hr of growth (average OD_{750} values for *Synechocystis* and UTEX 2973 cultures were 3.5 ± 0.2 and 3.6 ± 0.2, respectively). See **Supplementary Figure S2** for more info.

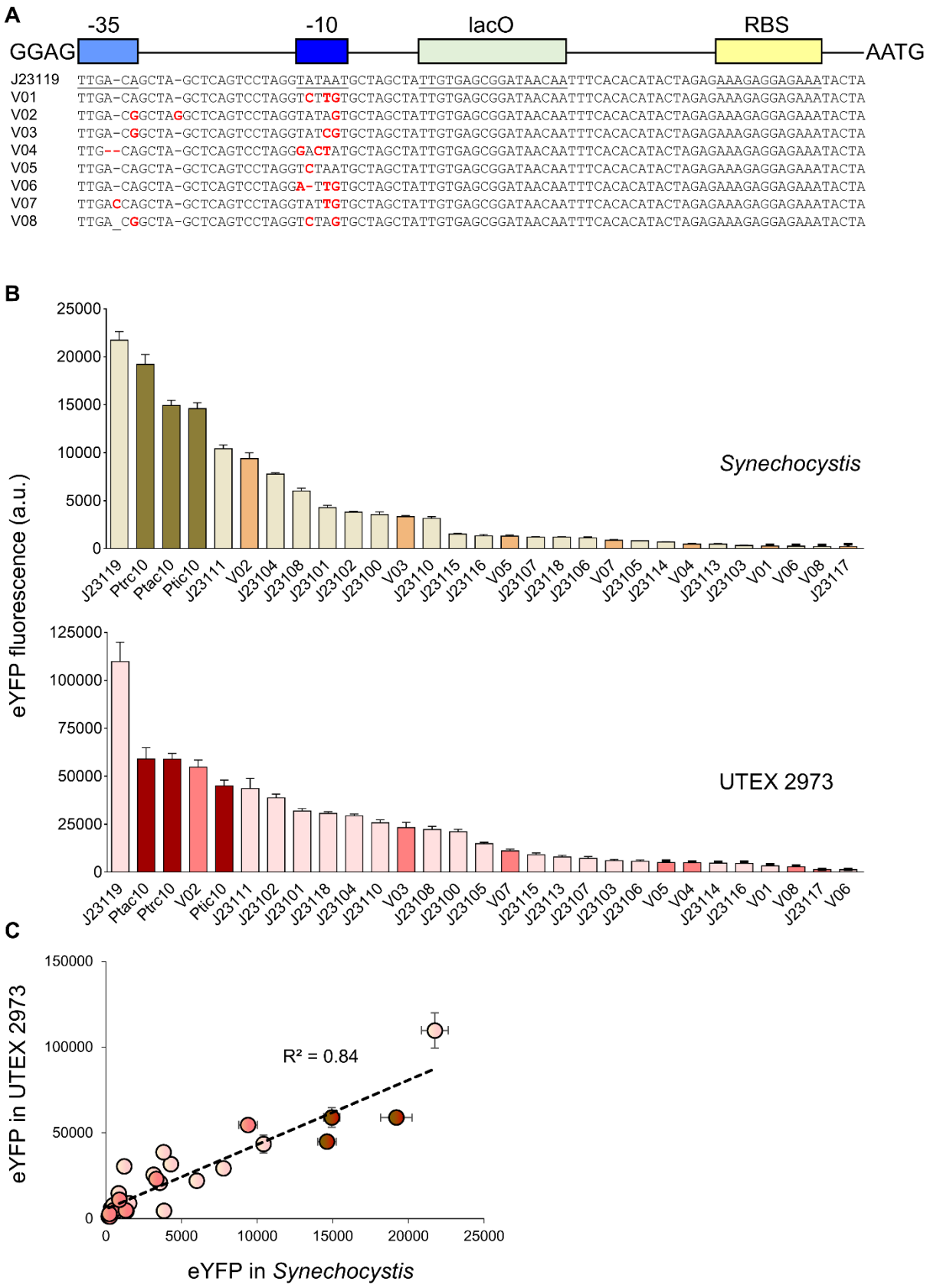


Figure 6. Expression levels of heterologous and synthetic promoters in *Synechocystis* and UTEX 2973. (A) Structure and alignment of eight new synthetic promoters derived from the BioBricks BBa_J23119 library and P_{trc10} promoter design (18). (B) Expression levels of eYFP

driven by promoters in *Synechocystis* and UTEX 2973 calculated from measurements taken from 10,000 individual cells. (C) Correlation analysis of expression levels of synthetic promoters tested in *Synechocystis* and UTEX 2973. The coefficient of determination (R^2) is shown for the J23119 library (red), new synthetic promoters (pink) and *trc* variants (dark red). Values are the means ± SE from at least four biological replicates after 48 hr of growth (average OD_{750} values for *Synechocystis* and UTEX 2973 cultures were 3.5 ± 0.2 and 3.6 ± 0.2, respectively). See **Supplementary Figure S2** for more info.

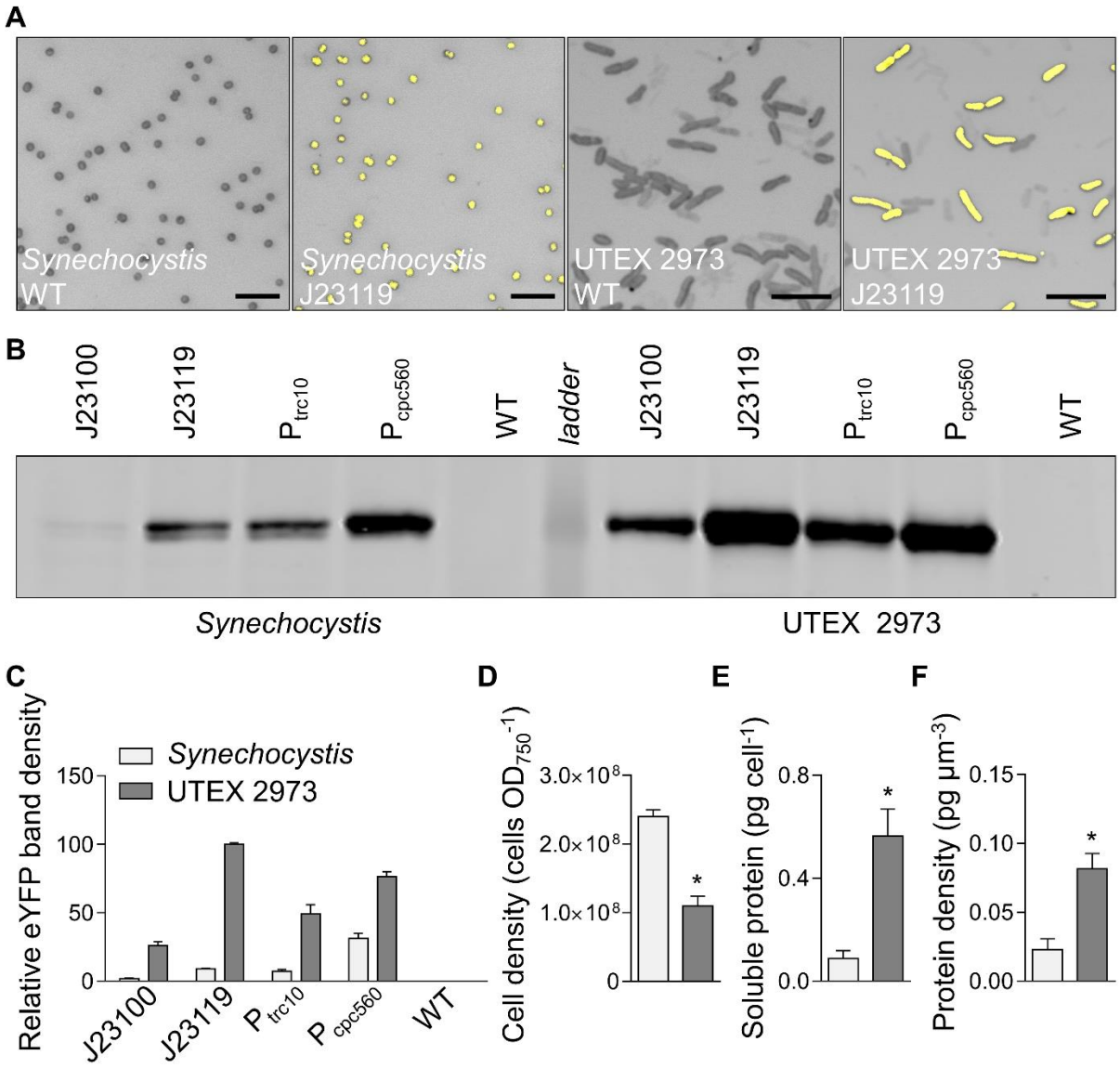


Figure 7. Protein expression levels in *Synechocystis* and UTEX 2973 cells. (A) Confocal images of WT strains and mutants expressing eYFP (fluorescence shown in yellow) driven by the J23119 promoter (bar = 10 µm). (B) Representative immunoblot of protein extracts (3 µg protein) from mutants with different promoter expression cassettes (as in **Fig. 6**) probed with an antibody against eYFP. The protein ladder band corresponds to 30 kDa. (E) Relative eYFP protein abundance relative to UTEX 2973 mutants carrying the J23119 expression cassette. (C-E) Cell density, protein content per cell and protein density per estimated cell volume for *Synechocystis* and UTEX 2973. Asterisks (*) indicate significant difference ($P < 0.05$) as determined by Student's *t*-tests. Values are the means ± SE of four biological replicates.

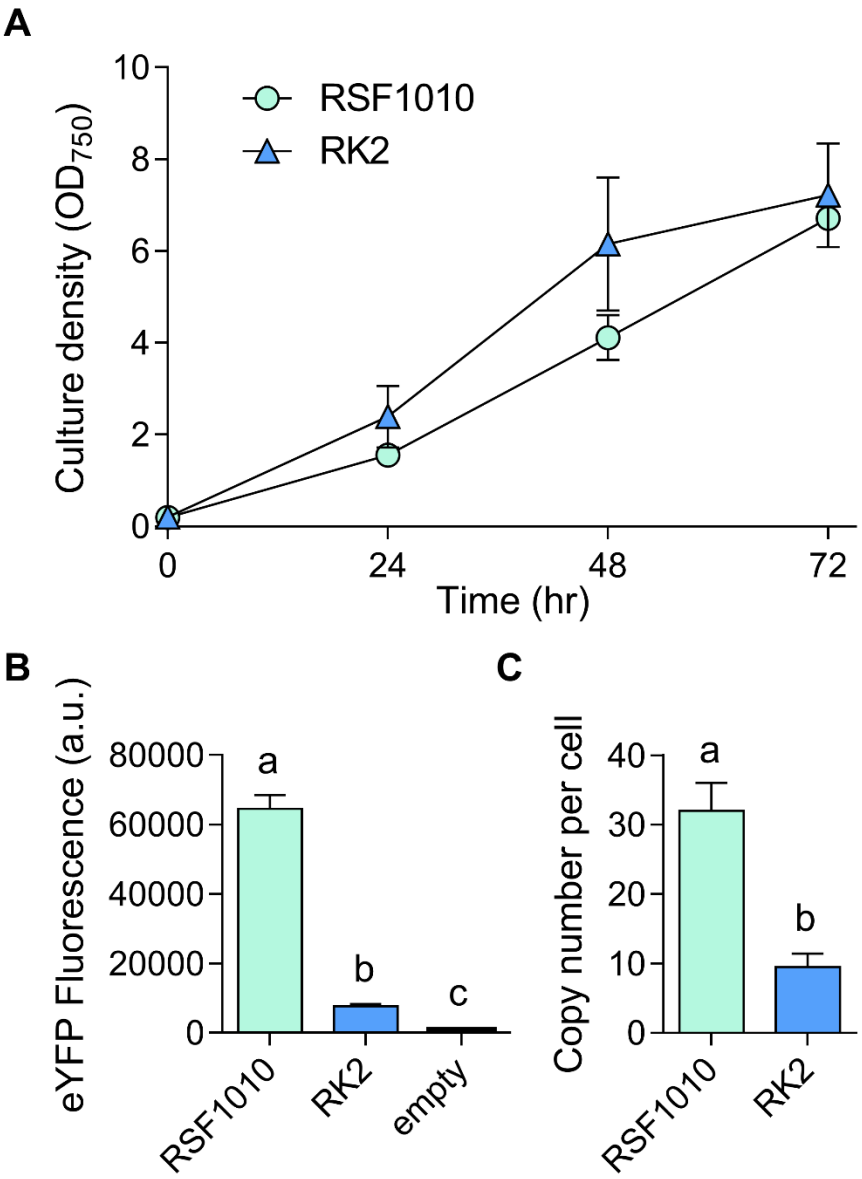


Figure 8. Growth and expression levels of eYFP with the RK2 replicative origin in *Synechocystis*. (A) Growth of strains carrying RK2 (vector pSEVA421-T-eYFP), RSF1010 (pPMQAK1-T-eYFP) or an empty pPMQAK1-T grown in appropriate antibiotics. Growth was measured as OD_{750} under a constant illumination of 100 µmol photons $m^{-2}s^{-1}$ at 30 °C. (B) Expression levels of eYFP after 48 hr of growth calculated from measurements taken from 10,000 individual cells. (C) Plasmid copy numbers per cell after 48 hr of growth. Letters indicating significant difference ($P < 0.05$) are shown, as determined by ANOVA followed by Tukey's HSD tests. Values are the means ± SE of four biological replicates.

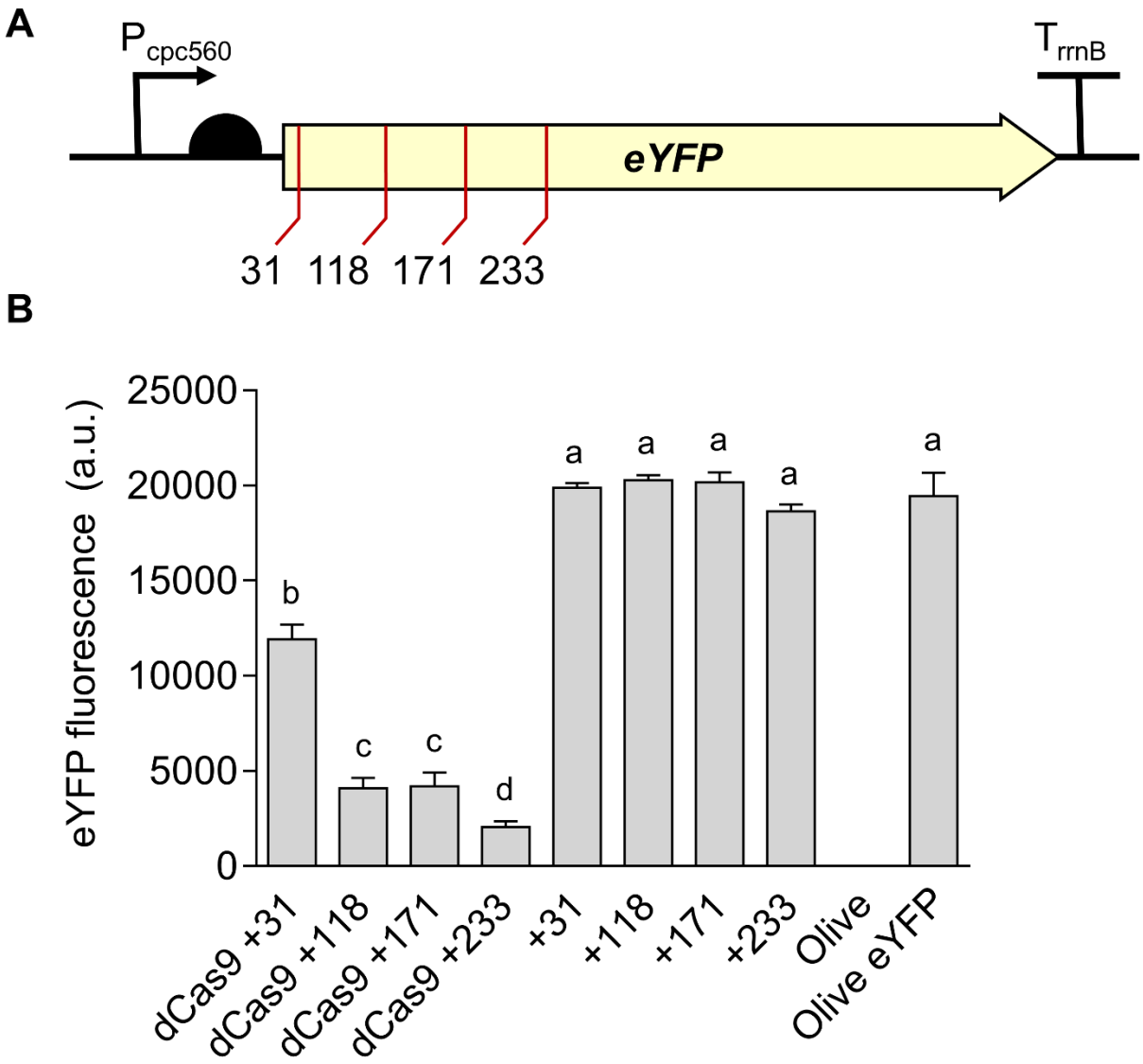


Figure 9. Gene regulation system using CRISPRi in *Synechocystis*. (A) Four target regions were chosen as sgRNA protospacers to repress eYFP expression in Olive-eYFP (**Fig. 4**): 'CCAGGATGGGCACCACCC' (+31), 'ACTTCAGGGTCAGCTTGCCGT' (+118), 'AGGTGGTCACGAGGGTGGGCCA' (+171) and 'AGAAGTCGTGCTGCTTCATG' (+233). (B) eYFP fluorescence of Olive-eYFP expressing constructs carrying sgRNAs with and without dCas9 (representative of 10,000 individual cells). Untransformed Olive-eYFP and the Olive mutant were used as controls. Letters indicating significant difference ($P < 0.05$) are shown, as determined by ANOVA followed by Tukey's HSD tests. Values are the means ± SE of four biological replicates.

SUPPLEMENTARY DATA

Supplementary Table S1. Table of all parts from CyanoGate kit generated in this work. Domestication refers to the removal of *BsaI* and/or *BsiI* sites (modifications are indicated in sequence maps provided in **Supp. Info. 2**). See separate .xlsx.

Supplementary Table S2. List of level T vectors used in this study.

Supplementary Table S3. Sequences of synthetic oligonucleotides used to determine copy number. Primers used for amplifying the *petB* locus were from Pinto et al. (2012).

Supplementary Information S1. Sequence maps (.gb files) of the components of the CyanoGate kit. See .zip file.

Supplementary Information S2. Protocols for MoClo assembly in level -1 through to level T. Protocols for assembly in level 0, level M and level T acceptor vectors (restriction enzyme *Bpi*I required, left). Protocols for assembly in level -1, level 1 and level P backbone vectors (restriction enzyme *Bsa*I required, right). Adapted from "A quick guide to Type IIS cloning" (Patron Lab; patronlab.org). For troubleshooting Type IIS mediated assembly we recommend synbio.tsl.ac.uk/docs.

Supplementary Information S3. Detailed assembly strategies using the CyanoGate kit.

Supplementary Information S4. Integrative engineering strategies using the CyanoGate kit. (A) Marked mutants are generated using a level T marked knock out vector carrying DNA sequences flanking the target locus of the chromosome (~1 kb), an antibiotic resistance cassette (Ab^R) and a sucrose selection cassette (*sacB*) that produces the toxic compound levansucrase in the presence of sucrose (20). Several rounds of segregation are required to identify a marked mutant. (B) Marked mutants then can be unmarked with a level T unmarked knock out vector and selection on sucrose-containing agar plates. (C) Unmarked knock in mutants can also be generated from marked mutants using a level T unmarked knock in vector carrying a gene expression cassette (UP FLANK LINKER and DOWN FLANK LINKER are shown in pink and light green, respectively). (D) Alternatively marked knock in mutants can be engineered in a single step using a level T marked knock vector (Ab^R UP LINKER and DOWN LINKER are shown in blue and orange, respectively). See **Fig. 2** for abbreviations.

Supplementary Information S5. Comparison of Gibson Assembly and Golden Gate Assembly. (A) A comparison of Gibson Assembly and Golden Gate Assembly pathways for building the level T vector cpcBA-eYFP described in **Fig. 4**. (B) Advantages and disadvantages of Gibson Assembly and Golden Gate Assembly.

Supplementary Information S6. Protocol and online interface for building CyanoGate vector assemblies. A CyanoGate online vector assembly tool called Design and Build (DAB) from the Edinburgh Genome Foundry.

Supplementary Figure S1. Comparison of growth for *Synechocystis*, PCC 7942 and UTEX 2973 under different culturing conditions. Values are the means ± SE from at least five biological replicates from two independent experiments.

Supplementary Figure S2. Growth and expression levels of heterologous and synthetic promoters in *Synechocystis* and UTEX 2973. (A) *Synechocystis* and UTEX 2973 was cultured for 72 hr at 30°C with continuous light (100 µmol photons m^{-2} s^{-1}) and 40°C with 300 µmol photons m^{-2} s^{-1}, respectively (see **Fig. 6**). Expression levels of eYFP are shown at three time points (24, 48 and 72 hr after inoculation). Values are the means ± SE from at least four biological replicates where each replicate represents the median measurements of 10,000 cells

Supplementary Figure S3. Cell volume calculations for *Synechocystis* and UTEX 2973 from confocal microscopy images. (A) Example confocal images of *Synechocystis* (left) and UTEX 2973 (right) cells expressing eYFP driven by the J23119 promoter at 48 hr. Individual cells were selected and measured using Leica AF Lite software (Leica Microsystems). Top panel: eYFP fluorescence (green); middle panel: chlorophyll auto florescence (red); bottom panel: overlay of eYFP and chlorophyll signals (yellow). (B) Volume estimations based on confocal image data (n =50) (C) Mathematical formulas used for calculating cell volume based on the cell shapes of *Synechocystis* (coccus, spherical) and UTEX 2973 (bacillus, cylindrical).

Supplementary Table S2. List of level T vectors used in this study.

Vector ID	Part name	Acceptor backbone	Level 1 vectors	Selection	Notes
Assembled Level T vector (integrative)					
pCAT.336	pUC19A-T (cpcBA-M)	pUC19A-T	1	Amp^R	Generated a marked mutant in the cpcBA promoter and operon (Fig. 3).
pCAT.337	pUC19A-T (cpcBA-UM)		1		Generated an unmarked mutant in the cpcBA promoter and operon (Fig. 3).
pCAT.312	pUC19A-T (cpcBA-eYFP)		3		Introduced a eYFP expression cassette into the marked ΔcpcBA "Olive" mutant (Fig. 4).
Assembled Level T vector (replicative)					
pCAT.9	pSEVA431-T	pSEVA431 Level T	-	$Spec^R$	Level T Acceptor, pBBR1 replicative origin (50).
pCAT.13	pSEVA442-T	pSEVA442 Level T	-		Level T Acceptor, pRO1600/ColE1 replicative origin (50).
pCAT.163	pSEVA421-T (P_{cpc560}-eYFP-T_{rrnB})	pSEVA421-T	1		Level T assembly with eYFP expression cassette, RK2 replicative origin (Fig. 8) (50).
pCAT.214	pPMQAK1-T (J23100 MH-eYFP-T_{rrnB})	pPMQAK1-T	1	Amp^R, Kan^R	Level T assemblies with eYFP expression cassette (Fig. 5-7) (18), pPMQAK1-T replicative origin.
pCAT.235	pPMQAK1-T (J23101MH-eYFP-T_{rrnB})		1		
pCAT.236	pPMQAK1-T (J23102MH-eYFP-T_{rrnB})		1		
pCAT.237	pPMQAK1-T (J23103MH-eYFP-T_{rrnB})		1		
pCAT.238	pPMQAK1-T (J23104MH-eYFP-T_{rrnB})		1		
pCAT.239	pPMQAK1-T (J23105MH-eYFP-T_{rrnB})		1		
pCAT.240	pPMQAK1-T (J23106MH-eYFP-T_{rrnB})		1		
pCAT.241	pPMQAK1-T (J23107MH-eYFP-T_{rrnB})		1		

pCAT.242	pPMQAK1-T (J23108MH-eYFP-T_{rrnB})	1
pCAT.243	pPMQAK1-T (J23109MH-eYFP-T_{rrnB})	1
pCAT.244	pPMQAK1-T (J23110MH-eYFP-T_{rrnB})	1
pCAT.245	pPMQAK1-T (J23111MH-eYFP-T_{rrnB})	1
pCAT.193	pPMQAK1-T (J23113MH-eYFP-T_{rrnB})	1
pCAT.247	pPMQAK1-T (J23114MH-eYFP-T_{rrnB})	1
pCAT.248	pPMQAK1-T (J23115MH-eYFP-T_{rrnB})	1
pCAT.249	pPMQAK1-T (J23116MH-eYFP-T_{rrnB})	1
pCAT.250	pPMQAK1-T (J23117MH-eYFP-T_{rrnB})	1
pCAT.251	pPMQAK1-T (J23118MH-eYFP-T_{rrnB})	1
pCAT.252	pPMQAK1-T (BBa_J23119MH-eYFP-T_{rrnB})	1
pCAT.253	pPMQAK1-T (P_{psbA2L}-eYFP-T_{rrnB})	1
pCAT.254	pPMQAK1-T (P_{rnpB}-eYFP-T_{rrnB})	1
pCAT.262	pPMQAK1-T (P_{trc10}-eYFP-T_{rrnB})	1
pCAT.263	pPMQAK1-T (P_{tac10}-eYFP-T_{rrnB})	1
pCAT.264	pPMQAK1-T (P_{tic10}-eYFP-T_{rrnB})	1
pCAT.267	pPMQAK1-T (J23119MH_V01-eYFP-T_{rrnB})	1
pCAT.268	pPMQAK1-T (J23119MH_V02-eYFP-T_{rrnB})	1
pCAT.269	pPMQAK1-T (J23119MH_V03-eYFP-T_{rrnB})	1
pCAT.272	pPMQAK1-T (J23119MH_V04-eYFP-T_{rrnB})	1
pCAT.273	pPMQAK1-T (J23119MH_V05-eYFP-T_{rrnB})	1

pCAT.274	pPMQAK1-T (J23119MH_V06-eYFP-T_{rrnB})	pPMQAK1-T	1	AmpR, KanR	Level T assemblies for CRISPRi (Fig. 9) (81), pPMQAK1-T replicative origin.
pCAT.265	pPMQAK1-T (J23119MH_V07-eYFP-T_{rrnB})		1		
pCAT.266	pPMQAK1-T (J23119MH_V08-eYFP-T_{rrnB})		1		
pCAT.278	pPMQAK1-T (P_{cpc560}+A-eYFP-T_{rrnB})		1		
pCAT.279	pPMQAK1-T (P_{cpc560}-eYFP-T_{rrnB})		1		
pCAT.280	pPMQAK1-T ($P_{cpc560\text{-}Ux2}$-eYFP-T_{rrnB})		1		
pCAT.281	pPMQAK1-T ($P_{cpc560\text{-}Dx2}$-eYFP-T_{rrnB})		1		
pCAT.314	pPMQAK1-T (P_{cpc560}-dCas9-T_{rrnB}) + (P_{trc10_TSS}-sgRNA+31-sgRNA scaffold)		2		
pCAT.315	pPMQAK1-T (P_{cpc560}-dCas9-T_{rrnB}) + (P_{trc10_TSS}-sgRNA+118-sgRNA scaffold)		2		
pCAT.316	pPMQAK1-T (P_{cpc560}-dCas9-T_{rrnB}) + (P_{trc10_TSS}-sgRNA+171-sgRNA scaffold)		2		
pCAT.317	pPMQAK1-T (P_{cpc560}-dCas9-T_{rrnB}) + (P_{trc10_TSS}-sgRNA+233-sgRNA scaffold)		2		
pCAT.319	pPMQAK1-T (P_{trc10_TSS}-sgRNA+31-sgRNA scaffold)		1		
pCAT.320	pPMQAK1-T (P_{trc10_TSS}-sgRNA+118-sgRNA scaffold)		1		
pCAT.321	pPMQAK1-T (P_{trc10_TSS}-sgRNA+171-sgRNA scaffold)		1		
pCAT.322	pPMQAK1-T (P_{trc10_TSS}-sgRNA+233-sgRNA scaffold)		1		

Supplementary Table S3. Sequences of synthetic oligonucleotides used to determine copy number. Primers used for amplifying the *petB* locus were from Pinto et al. (2012).

Name	Locus	Amplicon length (bp)	Forward primer	Reverse primer
pPMQAK1-T (RSF1010)	-	245	AGTTAAGCCAGCCCCGACACC	TTGAGTGAGCTGATACCGCT
pSEVA421-T (RK2)	-	135	ACGACCAAGAAGCGAAAAACC	CCACGGCGCAATATCGAAC
petB locus	-	1000	ATAGTACGCTGATTATATGCGATTTTACGG	CATGTAAAGAATGTCGTTGGGCCA
petB	slr0342 (Chr:2647386-2650184)	179	CCTTCGCCTCTGTCCAATAC	TAGCATTACACCCACAACCC
secA locus	-	1000	CATAACCTTCTTGCTTATATTCAATCAAGGGA	AGCCAGGAAACGGAAGACTTAC
secA	sll0616 (Chr:2428010-2428678)	113	TTAAATCCAAACCTTCCAGCACCC	AACCTATTACTACGACATCCGTAAGC

Supplementary Information S2. Protocols for MoClo assembly in level -1 through to level T. Protocols for assembly in level 0, level M and level T acceptor vectors (restriction enzyme *Bpi*I required, left). Protocols for assembly in level -1, level 1 and level P backbone vectors (restriction enzyme *Bsa*I required, right). Adapted from "A quick guide to Type IIS cloning" (Patron Lab; patronlab.org). For troubleshooting Type IIS mediated assembly we recommend synbio.tsl.ac.uk/docs.

_Bpi_I protocol (in restriction buffer)

- 50-100 ng of acceptor vector.
- For each modular vector/part to insert, use a 2:1 ratio of insert: acceptor.
- 2 µl 10 mM ATP (not dATP).
- 2 µl Buffer G (Thermofisher).
- 2 µl BSA (10X).
- 10 units *Bpi*I (1 µl 10 U/µl *Bpi*I, ThermoFisher)
- 200 units T4 DNA ligase (1 µl 200U/µl, ThermoFisher)

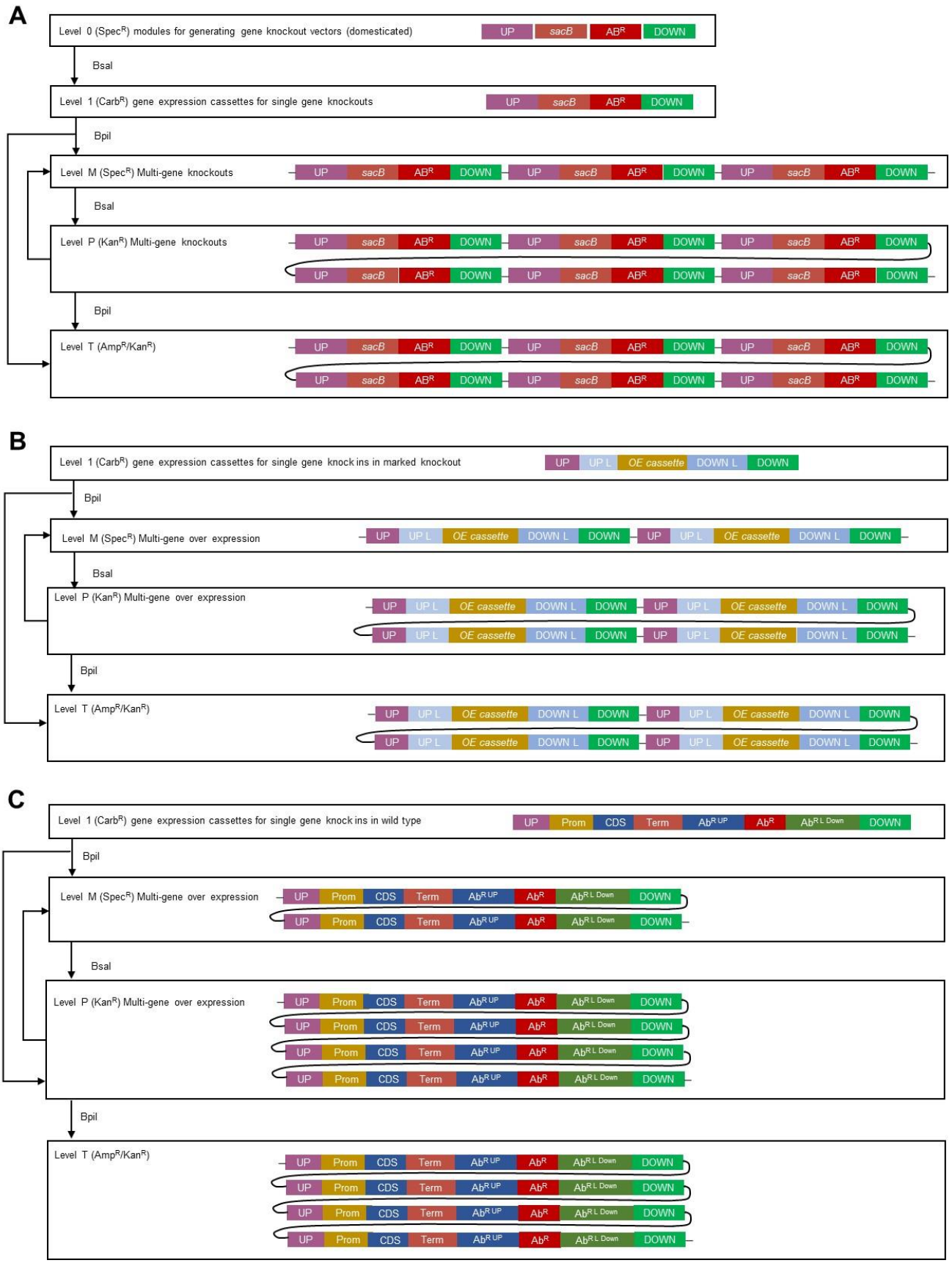
37^0 C for 10 minutes
16^0 C for 10 minutes
} X5
37^0 C for 20 minutes
65^0 C for 10 minutes
16^0 C (hold)

_Bsa_I protocol (in restriction buffer)

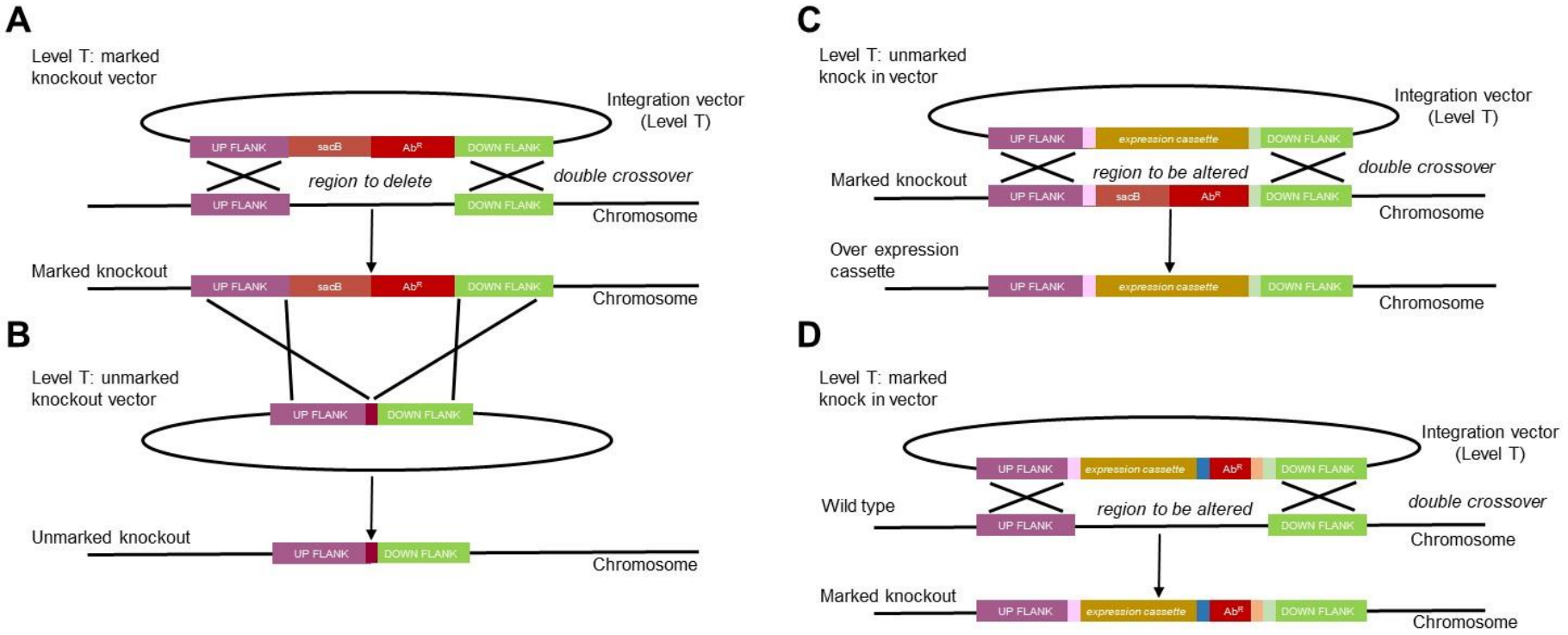
- 50-100 ng of acceptor vector.
- For each modular vector/part to insert, use a 2:1 ratio of insert: acceptor.
- 2 µl 10 mM ATP (not dATP).
- 2 µl Buffer G (Thermofisher).
- 2 µl BSA (10X).
- 10 units *Bsa*I (1 µl 10 U/µl *Bsa*I, ThermoFisher)
- 200 units T4 DNA ligase (1 µl 200U/µl, ThermoFisher)

37^0 C for 10 minutes
16^0 C for 10 minutes
} X5
37^0 C for 20 minutes
65^0 C for 10 minutes
16^0 C (hold)

Supplementary Information S3. Detailed assembly strategies using the CyanoGate kit.

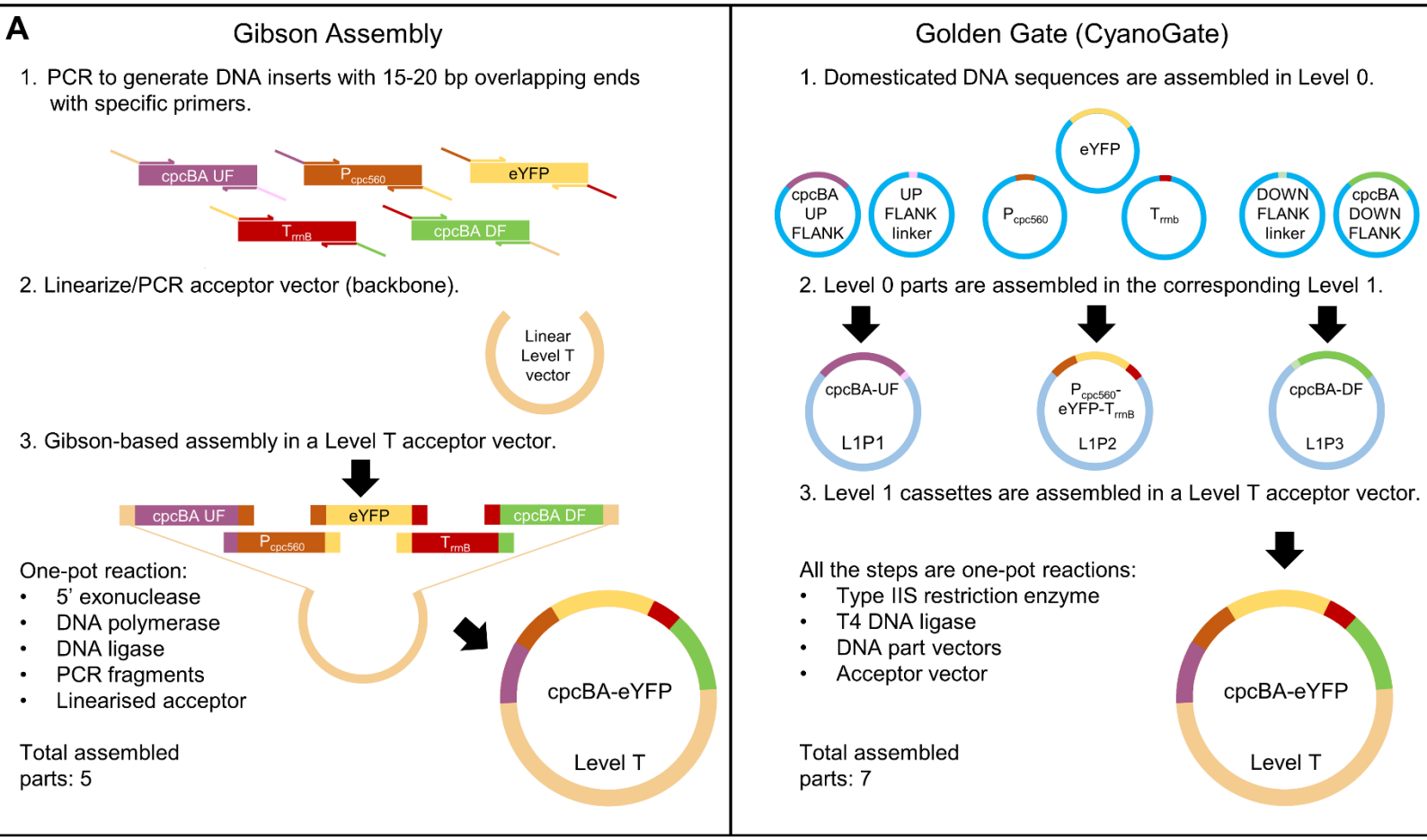


Supplementary Information S4. Integrative engineering strategies using the CyanoGate kit. (A) Marked mutants are generated using a level T marked knock out vector carrying DNA sequences flanking the target locus of the chromosome (~1 kb), an antibiotic resistance cassette (Ab^R) and a sucrose selection cassette (*sacB*) that produces the toxic compound levansucrase in the presence of sucrose (20). Several rounds of segregation are required to identify a marked mutant. (B) Marked mutants then can be unmarked with a level T unmarked knock out vector and selection on sucrose-containing agar plates. (C) Unmarked knock in mutants can also be generated from marked mutants using a level T unmarked knock in vector carrying a gene expression cassette (UP FLANK LINKER and DOWN FLANK LINKER are shown in pink and light green, respectively). (D) Alternatively marked knock in mutants can be engineered in a single step using a level T marked knock vector (Ab^R UP LINKER and DOWN LINKER are shown in blue and orange, respectively). See **Fig. 2** for abbreviations.



Supplementary Information S5. Comparison of Gibson Assembly and Golden Gate Assembly. (A) A comparison of Gibson Assembly and Golden Gate Assembly pathways for building the level T vector cpcBA-eYFP described in **Fig. 4**. (B) Advantages and disadvantages of Gibson Assembly and Golden Gate Assembly.

A



B

Description	**The Gibson Assembly (GA)** approach allows for the joining of two or more DNA fragments to generate plasmid vectors in a single isothermal reaction (Gibson et al., 2009).	**The Golden Gate (GG) Assembly** approaches (e.g. MoClo and GoldenBraid) use Type IIS restriction enzymes (REs) to generate standardised, non-palindromic overhangs that enable ordered assembly of multiple DNA parts in a single digestion-ligation reaction (Vazquez-Vilar et al., 2018).
Advantages	• Virtually any DNA fragments and any plasmid can be assembled together without prior modifications. • Allows seamless (scarless), directional cloning of multiple DNA fragments. • Can be used for cloning a wide range of DNA fragment sizes (i.e. 100-100,000 bp).	• Can re-use parts without modification in new assemblies. • Once parts are made, no subsequent PCR or clean-ups steps are required, and new assemblies do not require sequence checking. • GG only requires liquid handing (no columns or gels) so can be automated. Thus, GG is simple to scale for high-throughput protocols (e.g. assembly of combinatorial libraries).

Advantages	• Depending on the number of fragments, GA can help to avoid multiple rounds of cloning (i.e. into different levels). • GA does not require DNA domestication (i.e. removal of incompatible restriction enzyme recognition sites). • Apart from vector assembly, GA can be used for numerous additional applications such, site-directed mutagenesis, library construction, shotgun cloning and the development of bacterial artificial chromosomes (BACs) (Li et al. 2018).	• GG allows for the standardisation of parts and vectors:- o Standard overhangs allow for directional and hierarchical assembly. o Assemblies are carried out with a common set of established acceptor vectors and a defined assembly protocol. o Standard antibiotic selection markers and visual colony screening (e.g. blue/white) at each assembly level to facilitate the detection of positive colonies. o Establishment of a common genetic syntax (i.e. the Phytobricks standard) has enabled broader exchange of parts and assemblies (Patron et al., 2015). o GG simplifies experimental replication, and comparable information is available for part performance and methods for reliable assembly. o The availability of libraries of standard exchangeable DNA parts (e.g. Phytobricks, MoClo).
Disadvantages	• Primers for each part are needed for every assembly. • Unique overlapping primer pairs are required to join two different DNA fragments. This can limit the ability to freely combine different parts (e.g. for promoter screening). • PCR can fail. • Secondary structures and/or repetitive sequences in the overlap region can limit the efficiency and accuracy of assembly. • Sequence verification of all regions that undergo PCR amplification is recommended. Some DNA regions are challenging to sequence (e.g. the pPMQAK1 backbone), which can increase the cost of sequencing. • Assembly efficiencies decline with six or more DNA fragments or with the use of fragments shorter than 100 bp. • • Very small parts have to be first assembled by extension/overlap PCR.	• DNA parts and acceptor vectors require domestication to remove illegal Type IIS RE sites. • Some DNA sequences (e.g. promoters) may be challenging to domesticate due to the presence of RE sites in regulatory elements. • Assembly is quasi-seamless due to the use of standardised overhangs. • Initial setup can be time consuming, and purchase of Addgene kits could be a relatively costly starting investment.

Supplementary Information S6. Protocol and online interface for building CyanoGate vector assemblies. A CyanoGate online vector assembly tool called Design and Build (DAB) from the Edinburgh Genome Foundry.

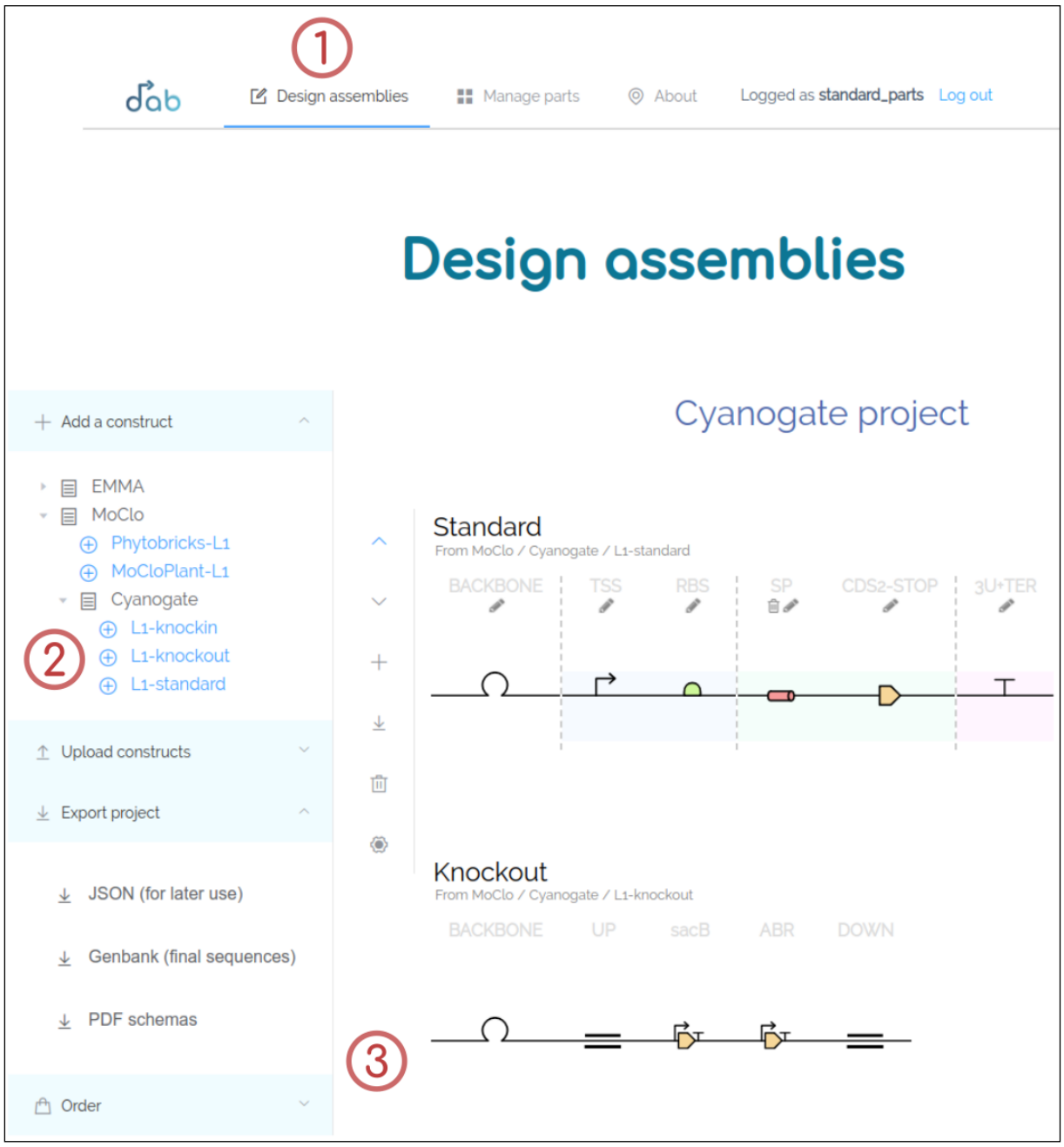
1. Site: dab.genomefoundry.org
2. Select "Home" and "Design New Assemblies".
3. Select "MoClo", then from the drop down list select "CyanoGate".
4. There are 3 options: a) L1-knockout, b) L1-knock in, and c) L1-standard.

   a. L1-knockout
      - This is a level 1 assembly for generating a marked or unmarked knockout mutant (see **Fig. 2**). The level 1 module(s) then should be transferred to the integrative level T integrative vector (pCAT15.UC19) for chromosomal integration in species amenable to natural transformation (e.g. *Synechocystis* sp. PCC 6803) (20).
      - Example of level 1 assembly for generating a level T knockout vector: L1P1 acceptor (DOWN FLANK + *sacB* + Ab^{R}Kan + UP FLANK).

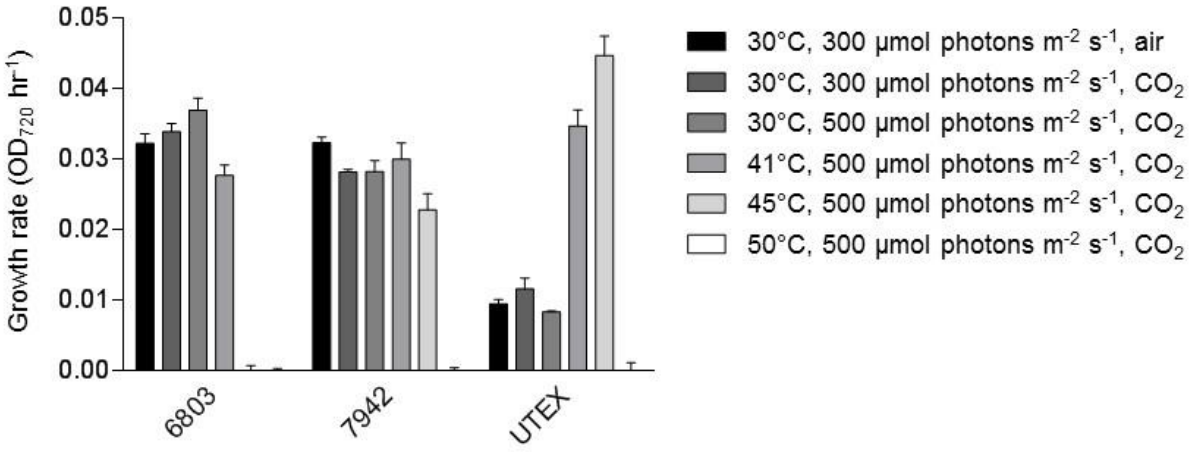
   b. L1-knock in
      - This is a level 1 assembly for generating a knock in mutant (see **Fig. 3**). Each level 0 flanking region should be assembled into a specific level 1 position with gene expression cassettes (L1-standard) in between them.
      - The level 1 modules then should be transferred to the integrative level T integrative vector (pCAT15.UC19) for chromosomal integration in species amenable to natural transformation (e.g. *Synechocystis* sp. PCC 6803).
      - Example of 3 level 1 assemblies for generating a level T knock in vector: L1P1 acceptor (6803 NS1 Down Flank (slr0573) + DOWN FLANK), L1P2 acceptor (P_{trc10} + eYFP + T_{rrnB}), L1P3 acceptor (UP FLANK + 6803 NS1 Up Flank (slr0573)).

   c. L1-standard.
      - This is a level 1 assembly for generating a standard gene expression cassette from level 0 parts.
      - These level 1 modules can be transferred to the integrative level T integrative vector (pCAT15.UC19) for chromosomal integration in species amenable to natural transformation (e.g. *Synechocystis* sp. PCC 6803).
      - Alternatively, the level 1 modules can be transferred to a replicative level T vector (e.g. pCAT0.PMQAK1) for transformation into cyanobacterial species amenable to conjugation or electroporation.
      - Example of a level 1 assembly for generating a level T expression vector: L1P1 acceptor (P_{trc10} + eYFP + T_{rrnB}).

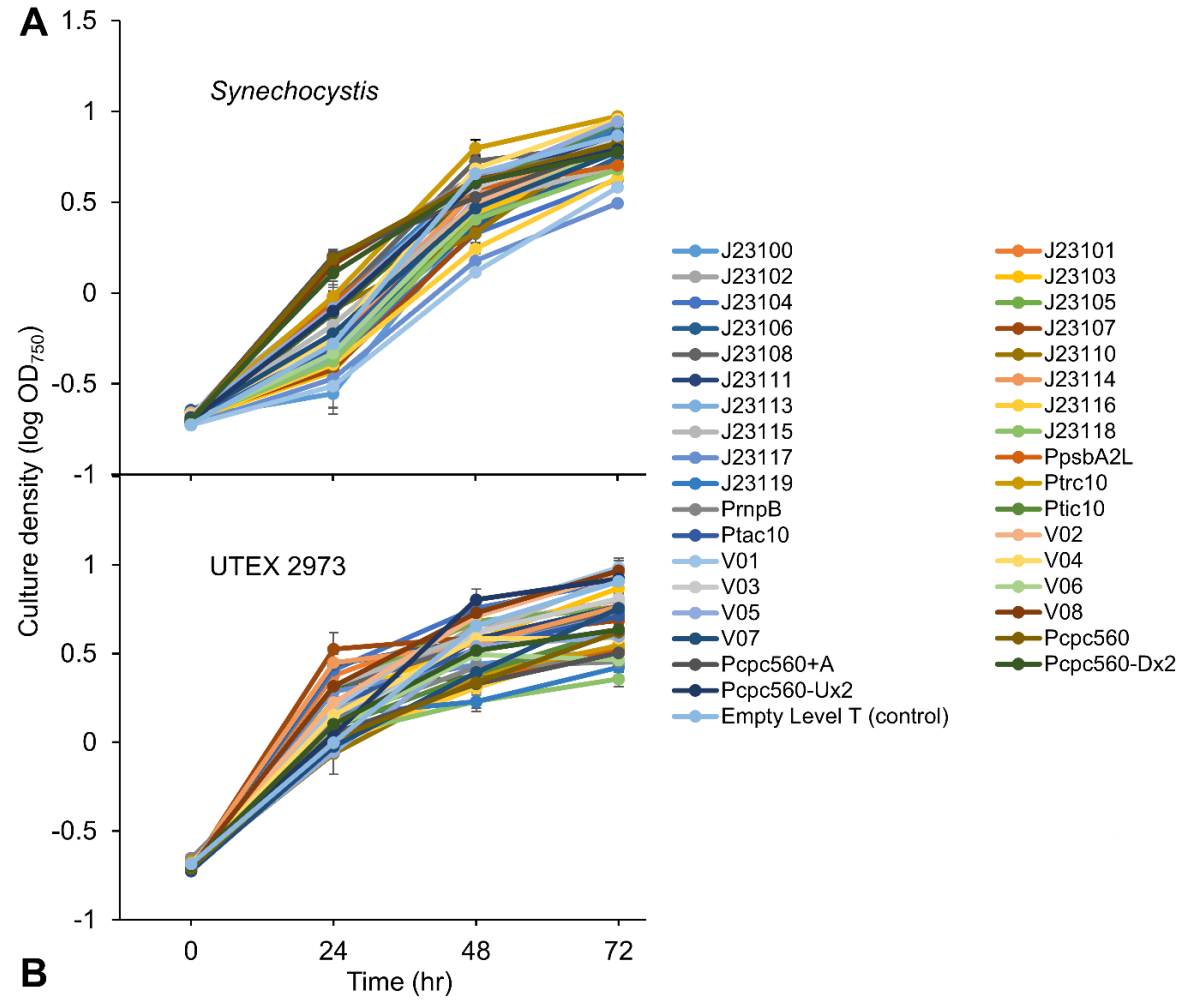
Figure. 1. Screenshot of the online "Design and Build" (DAB) tool (dab.genomefoundry.org) that allows users to browse parts and create structurally valid vector assemblies (1), choose from pre-defined templates (2), and order *in silico* assemblies directly from the foundry (3).



Supplementary Figure S1. Comparison of growth for *Synechocystis*, PCC 7942 and UTEX 2973 under different culturing conditions. Values are the means ± SE from at least five biological replicates from two independent experiments.

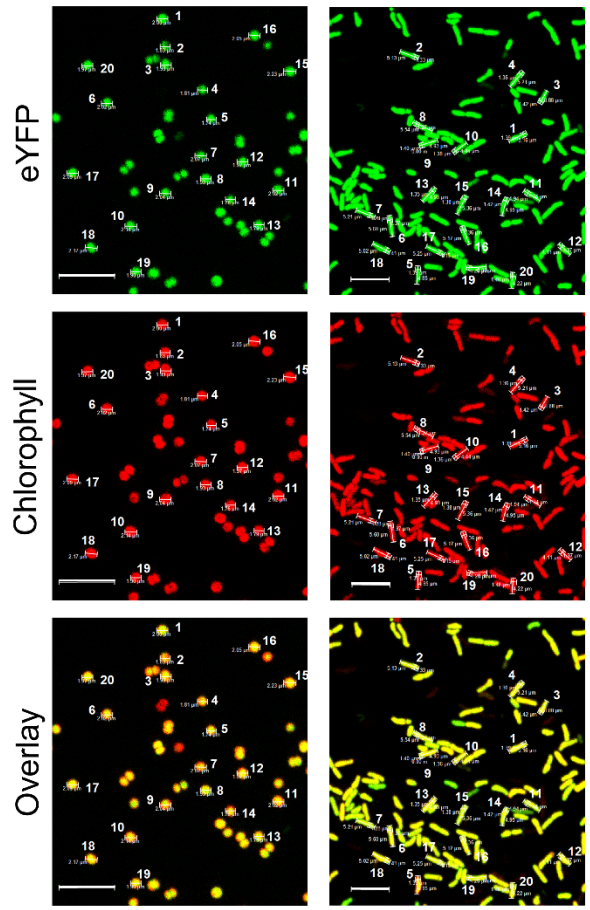


Supplementary Figure S2. Growth and expression levels of heterologous and synthetic promoters in *Synechocystis* and UTEX 2973. (A) *Synechocystis* and UTEX 2973 was cultured for 72 hr at 30°C with continuous light (100 µmol photons m^{-2} s^{-1}) and 40°C with 300 µmol photons m^{-2} s^{-1}, respectively (see **Fig 6**). Expression levels of eYFP are shown at three time points (24, 48 and 72 hr after inoculation). Values are the means ± SE from at least four biological replicates where each replicate represents the median measurements of 10,000 cells.


A
Synechocystis
UTEX 2973
Culture density (log OD750)
Time (hr)
J23100
J23101
J23102
J23103
J23104
J23105
J23106
J23107
J23108
J23110
J23111
J23114
J23113
J23116
J23115
J23118
J23117
PpsbA2L
J23119
Ptrc10
PrnpB
Ptic10
Ptac10
V02
V01
V04
V03
V06
V05
V08
V07
Pcpc560
Pcpc560+A
Pcpc560-Dx2
Pcpc560-Ux2
Empty Level T (control)
B
24 hr
48 hr
72 hr
Synechocystis
UTEX 2973
eYFP fluorescence (a.u.)

Supplementary Figure S3. Cell volume calculations for *Synechocystis* and UTEX 2973 from confocal microscopy images. (A) Example confocal images of *Synechocystis* (left) and UTEX 2973 (right) cells expressing eYFP driven by the J23119 promoter at 48 hr. Individual cells were selected and measured using Leica AF Lite software (Leica Microsystems). Top panel: eYFP fluorescence (green); middle panel: chlorophyll autoflorescence (red); bottom panel: overlay of eYFP and chlorophyll signals (yellow). (B) Volume estimations based on confocal image data (n =50) (C) Mathematical formulas used for calculating cell volume based on the cell shapes of *Synechocystis* (coccus, spherical) and UTEX 2973 (bacillus, cylindrical).

A



B

#Cell	UTEX 2973			*Synechocytis*	
	L	W	Volume	W	Volume
1	5.17	1.39	7.845	2.0	4.189
2	5.13	1.33	7.127	1.83	3.209
3	3.88	1.42	6.145	1.93	3.764
4	5.21	1.36	7.568	1.81	3.105
5	4.85	1.37	7.149	1.74	2.758
6	5.6	1.37	8.255	2.02	4.316
7	5.21	1.28	6.704	2.07	4.644
8	5.54	1.27	7.018	1.93	3.764
9	4.93	1.4	7.589	2.04	4.445
10	4.44	1.36	6.450	2.14	5.131
11	4.94	1.34	6.967	2.02	4.316
12	4.11	1.37	6.059	1.97	4.003
13	4.93	1.35	7.057	1.79	3.003
14	4.65	1.42	7.364	1.76	2.855
15	5.36	1.38	8.017	2.23	5.806
16	5.17	1.36	7.510	2.05	4.511
17	5.25	1.15	5.453	2.09	4.780
18	5.02	1.41	7.838	2.17	5.350
19	5.3	1.26	6.609	1.9	3.591
20	4.22	1.48	7.260	1.97	4.003
21	4.96	1.55	9.359	2.03	4.380
22	4.69	1.34	6.614	1.96	3.942
23	5.22	1.35	7.472	1.9	3.591
24	4.01	1.31	5.405	1.82	3.157
25	5.01	1.3	6.650	1.97	4.003
26	4.7	1.5	8.306	2	4.189
27	4.78	1.35	6.842	1.91	3.648
28	4.65	1.31	6.267	1.97	4.003
29	5.41	1.23	6.428	1.88	3.479
30	4.06	1.27	5.143	2.2	5.575
31	5.7	1.37	8.402	1.85	3.315
32	4	1.29	5.228	1.93	3.764
33	4.29	1.21	4.933	2.11	4.919
34	4.04	1.46	6.764	1.83	3.209
35	4.05	1.4	6.235	1.95	3.882
36	5.24	1.28	6.743	2.16	5.277
37	4.76	1.38	7.120	1.88	3.479
38	5.03	1.48	8.653	1.85	3.315
39	4.58	1.54	8.531	1.78	2.953
40	4.18	1.29	5.463	1.99	4.126
41	5.23	1.25	6.418	1.79	3.003
42	5.3	1.5	9.366	1.96	3.942
43	4.29	1.19	4.771	1.57	2.026
44	3.91	1.34	5.514	1.98	4.064
45	3.99	1.42	6.319	1.95	3.882
46	4.58	1.25	5.621	1.97	4.003
47	4.75	1.55	8.963	1.91	3.648
48	4.84	1.22	5.658	1.89	3.535
49	3.59	1.24	4.335	1.94	3.823
50	4.29	1.37	6.324	1.98	4.064
Average	**4.74**	**1.35**	**6.837**	**1.97**	**3.915**
stdev	**0.52**	**0.09**	**1.177**	**0.13**	**0.748**
sterror	**0.07**	**0.01**	**0.166**	**0.02**	**0.106**

C

Strain	*Synechocystis*	UTEX 2973
Shape	Sphere	Cylinder
Formula	$V = 4/3\,\pi r^3$	$V = \pi r^2 h$